

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

+

1/78

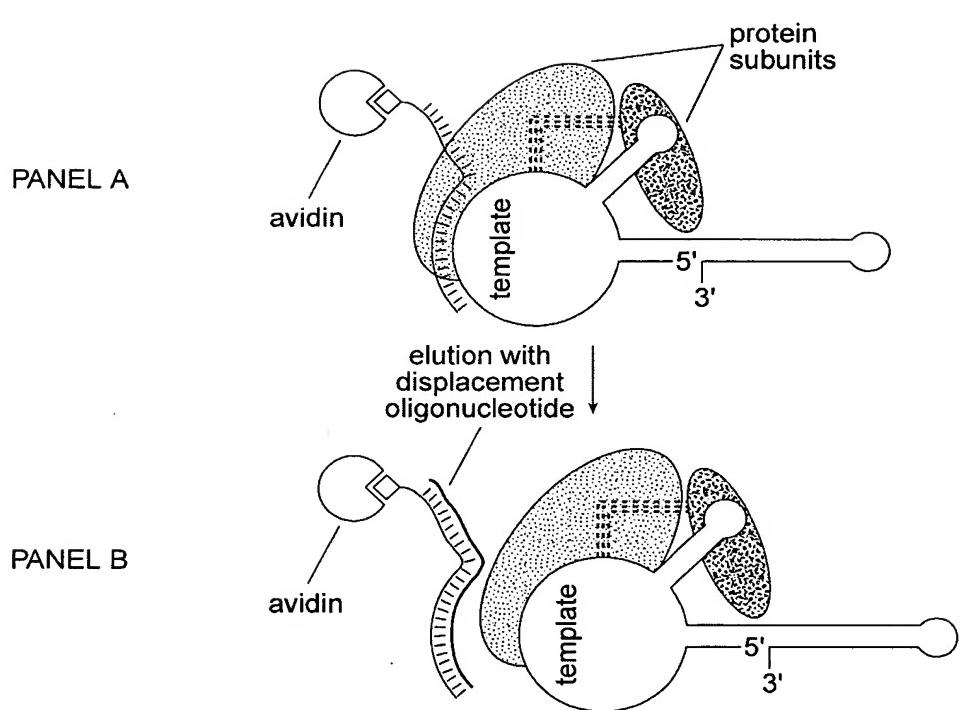


FIG. 1

+

+

2/78

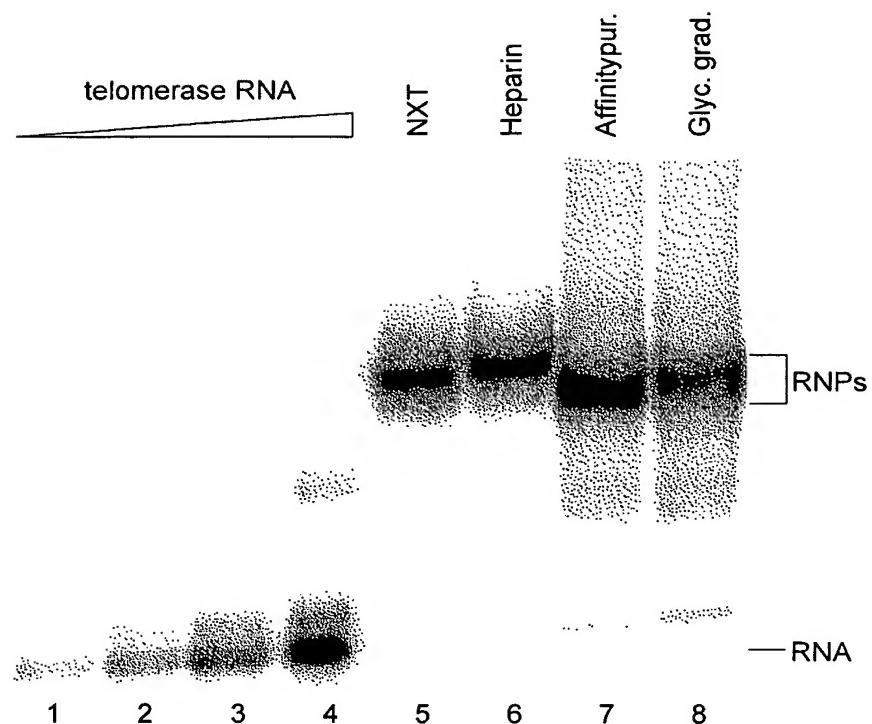


FIG. 2

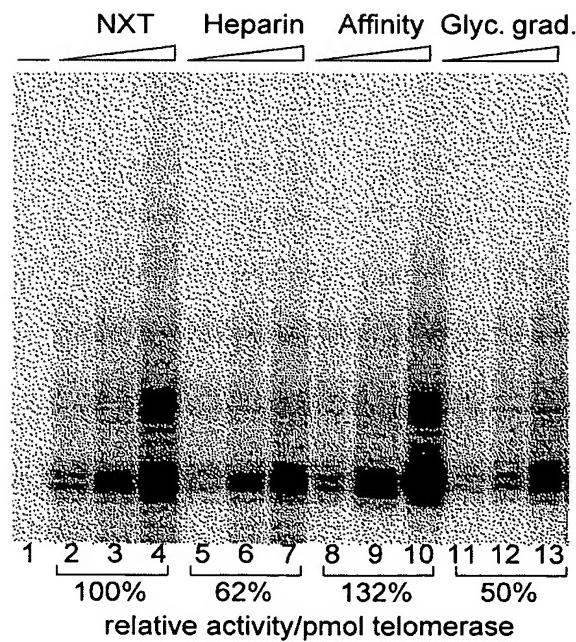


FIG. 3

+

+

3/78

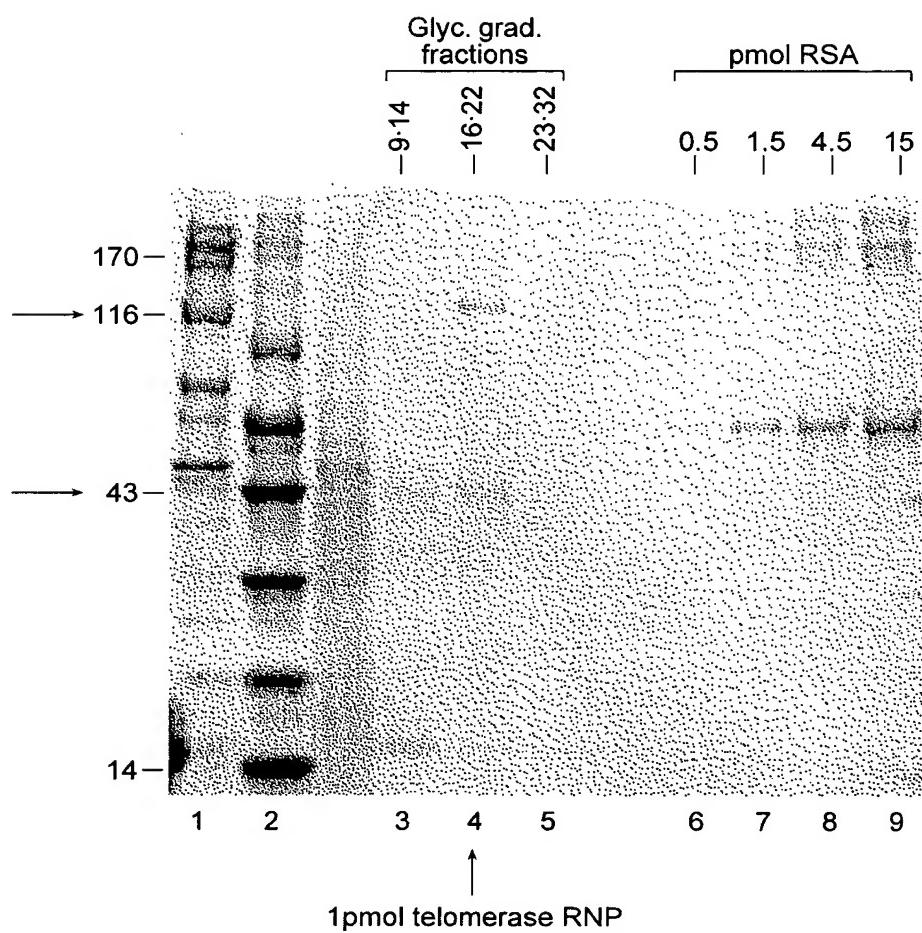


FIG. 4

+

+

4/78

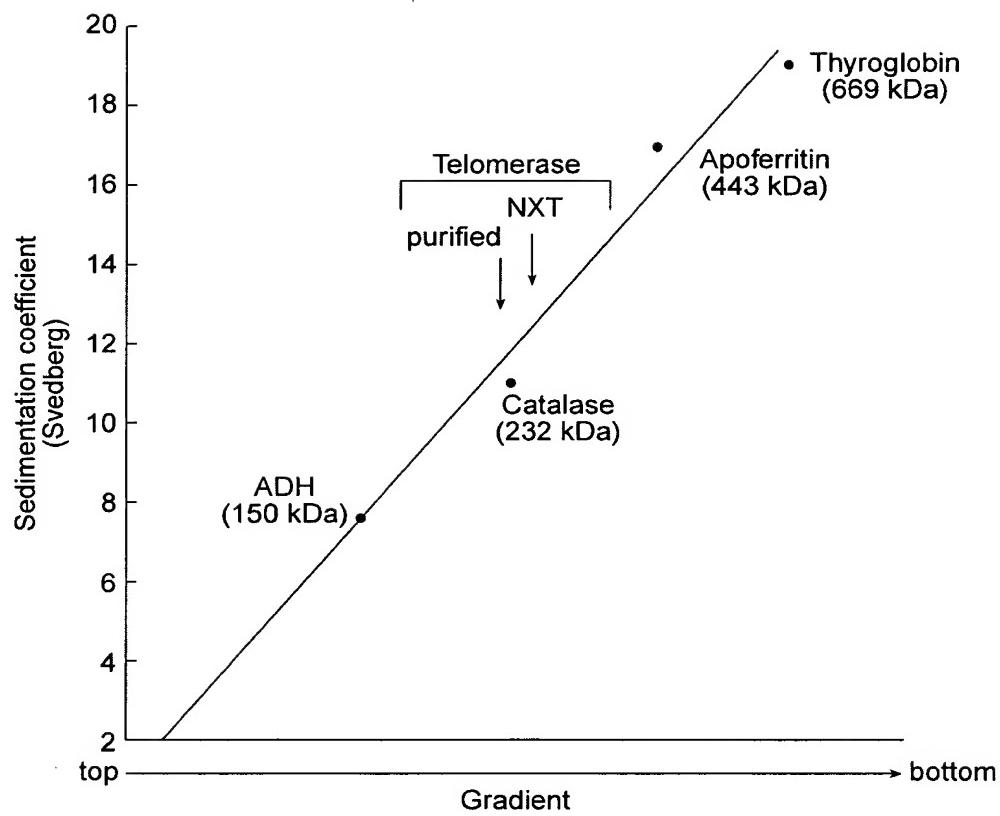


FIG. 5

+

O I P E
JUL 26 2004
PATENT & TRADEMARK OFFICE
911

5/78

Telomerase:

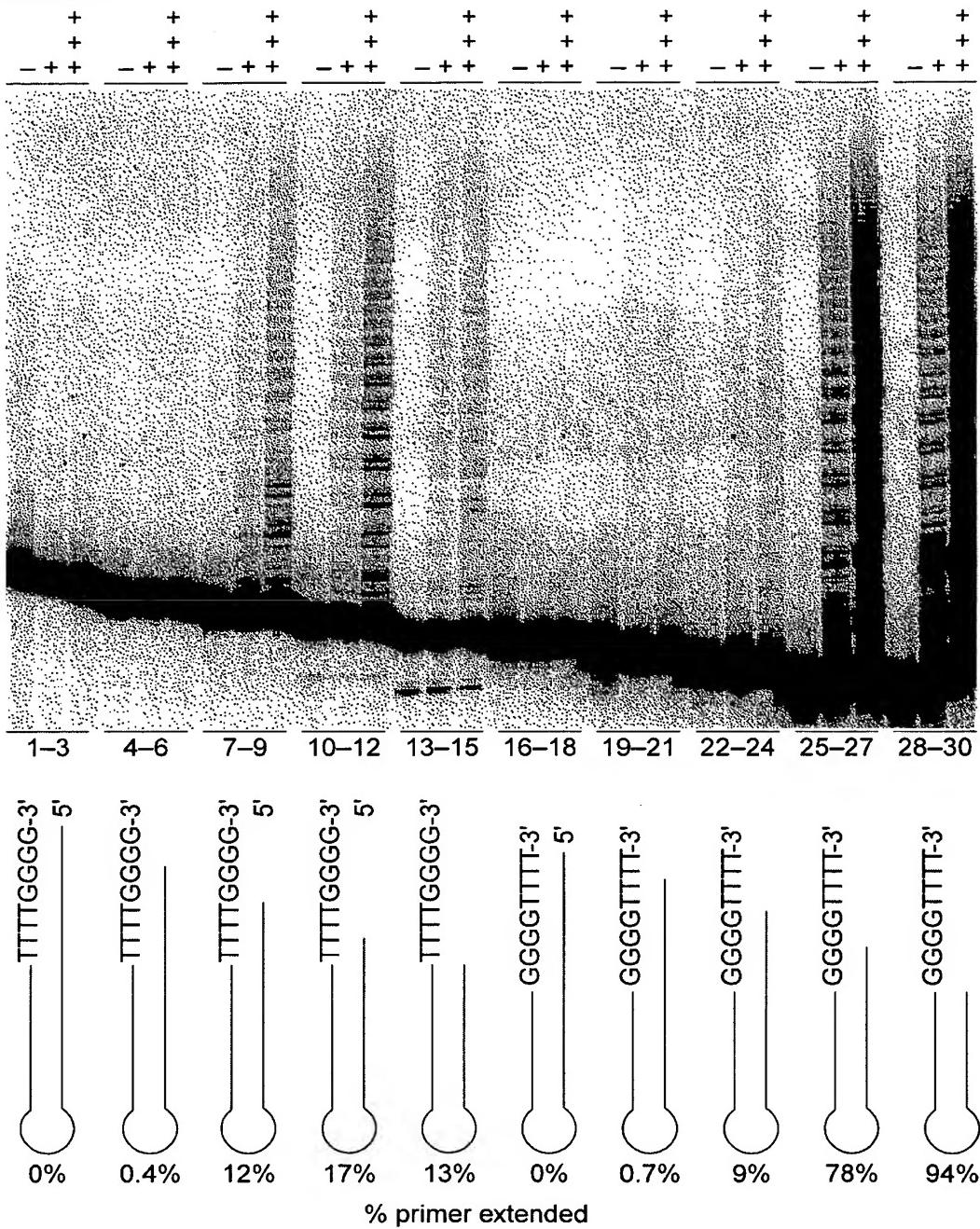


FIG. 6



6/78

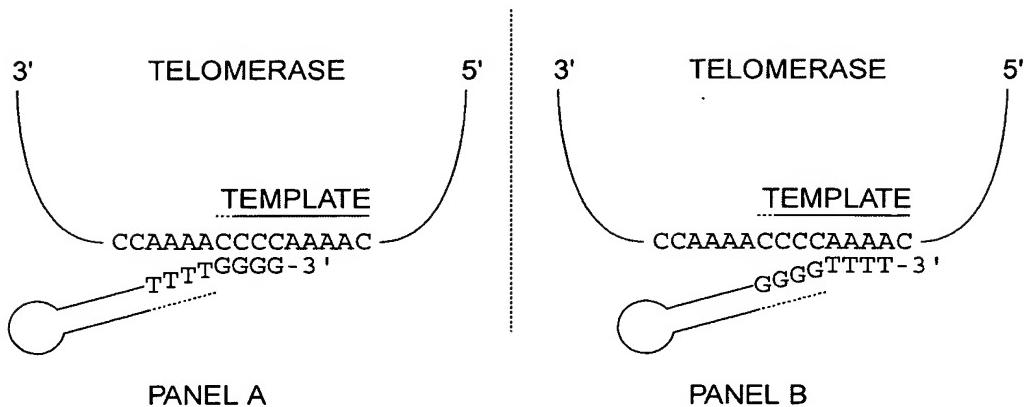


FIG. 7

1	CCCCAAAACC	CCAAAACCCC	AAAACCCCTA	TAAAAAAAGA	AAAAATTGAG
51	GTAGTTAGA	AATAAAATAT	TATTCCCGCA	CAAATGGAGA	TGGATATTGA
101	TTGGATGAT	ATAGAAAATT	TACTTCCTAA	TACATTCAAC	AAGTATAGCA
151	GCTCTTGTAG	TGACAAGAAA	GGATGCAAAA	CATTGAAATC	TGGCTCGAAA
201	TCGCCTTCAT	TGACTATTCC	AAAGTTGCAA	AAACAATTAG	AGTTCTACTT
251	CTCGGATGCA	AATCTTATA	ACGATTCTTT	CTTGAGAAAA	TTAGTTTAA
301	AAAGCGGAGA	GCAAAGAGTA	GAATTGAAA	CATTACTAAT	GTAAATAATAA
351	AATCAGGTAA	TGAGGATTAT	TCTATTTTTT	AGATCACTTC	TTAAGGAGCA
401	TTATGGAGAA	AATTACTTAA	TACTAAAGG	AAACAGTTT	GGATTATTTC
451	CCTAGCCAAC	AATGATGAGT	ATATTTAAATT	CATATGAGAA	TGAGTCAAAG
501	GATCTCGATA	CATCAGACTT	ACCAAAGACA	AACTCGCTAT	AAAACGCAAG
551	AAAAAGTTTG	ATAATCGAAC	AGCAGAAGAA	CTTATTGCA	TTACTATTG
601	TATGGGTTTT	ATTACAATTG	TTTAGGTAT	CGACGGTGAA	CTCCCGAGTC
651	TTGAGACAAT	TGAAAAAGCT	GTTTACAAC	GAAGGAATCG	CAGTTCTGAA
701	AGTCTGTGATG	TGTATGCCAT	TATTTGTGA	ATTAATCTCA	AATATCTTAT
751	CTCAATTAA	TGGATAGCTA	TAGAAACAAA	CCAAAATAAC	CATGCAAGTT
801	TAATGGAATA	TACGTTAAAT	CCTTGGGAC	AAATGCAACAC	TGAATTATA
851	TTGGATTCTT	AAAGCATAGA	TACACAGAAT	GCTTTAGAGA	CTGATTTAGC
901	TTACAACAGA	TTACCTGTTT	TGATTACTCT	TGCTCATCTC	TTATATCTTT
951	AAAAGAAGCA	GGCGAAATGA	AAAGAAGACT	AAAGAAAGAG	ATTTCAAAAT
1001	TTGTTGATTC	TTCTGTAACC	GGATTAAACA	ACAAGAATAT	TAGCAACGAA
1051	AAAGAAGAAG	AGCTATCAC	ATCTGTGATTC	TTAAAGATT	AAAAATTCC
1101	AGGTAAAGAGA	GATACATTCA	TTAAAATTCA	TATATTATAG	TTTTTCATTT
1151	CACAGCTGTT	ATTTTCTTT	ATCTTAAACAA	TATTTTTGAA	TTAGCTGGAA
1201	GTAAAAAGTA	TCAAATAAGA	GAAGCGCTAG	ACTGAGGTAA	CTTAGCTTAT
1251	TCACATTCA	AGATCGACCT	TCATATATCC	AATACGATGA	TAAGGAAACA
1301	GCAGTCATCC	GTTTTAAAAA	TAGTGCTATG	AGGACTAAAT	TTTTAGAGTC
1351	AAGAAATGGA	GCGAAATCT	TAATCAAAAAA	GAATTGCGTC	GATATTGCAA
1401	AAGAAATCGAA	CTCTAAATCT	TTCGTTAATA	AGTATTACCA	ATCTTGATTC
1451	ATTGAAGAGA	TTGACGAGGC	AACTGCACAG	AAGATCATTAA	AAGAAATAAA
1501	GTAACTTTA	TTAATTAGAG	AATAAAACTAA	ATTACTAATA	TAGAGATCAG
1551	CGATCTTCAA	TTGACGAAAT	AAAAGCTGAA	CTAAAGTTAG	ACAATAAAA
1601	ATACAAACCT	TGGTCAAAAT	ATTGAGGAAG	GAAAAGAAGA	CCAGTTAGCA
1651	AAAGAAAAAA	TAAGGCAATA	AATAAAATGA	GTACAGAACT	GAAGAAATAA
1701	AAGATTTATT	TTTTTCAATA	ATTTATTGAA	AAGAGGGTT	TTGGGGTTTT
1751	GGGGTTTTGG	GG			

FIG. 11



7/78

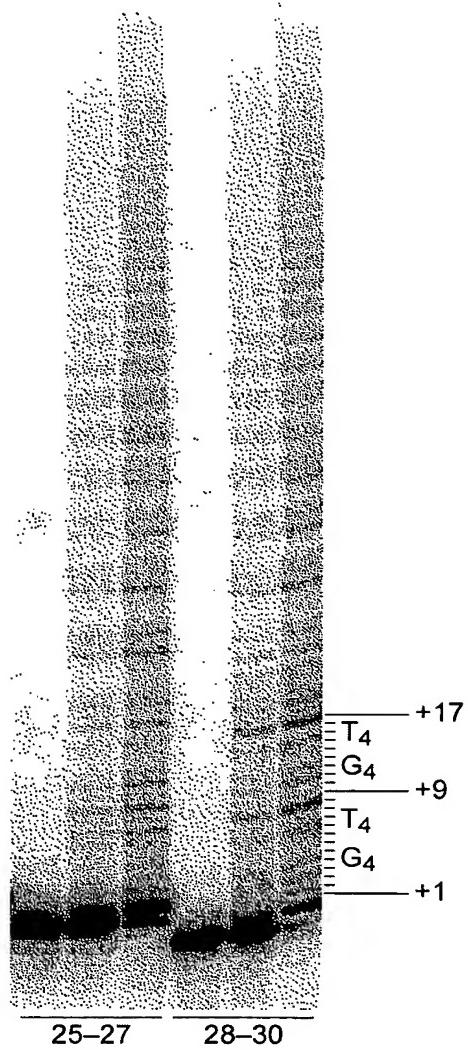


FIG. 8



8/78

1	AAAACCCCAA	AACCCCCAAA	CCCCTTTAG	AGCCCTGCAG	TTGGAAATAT
51	AACCTCAGTA	TTAATAAGCT	CAGATTTAA	ATATTAATTAA	CAAAACCTAA
101	ATGGAGGTTG	ATGTTGATAA	TCAAGCTGAT	AATCATGGCA	TTCACTCAGC
151	TCTTAAGACT	TGTGAAGAAA	TTAAAGAACG	AAAACGTTG	TAACCTTGGA
201	TCCAGAAAGT	TATTAGATGA	AGAAAATCAAT	CTCAAAGTCA	TTATAAAGAT
251	TTAGAAAGATA	TTAAAATATT	TGCGCAGACAA	AATATTGTTG	CTACTCCACG
301	AGACTATAAT	GAAGAAAGATT	TTAAAGTTAT	TGCAAGAAAA	GAAGTATTTT
351	CAACTGGACT	AATGATCGAA	CTTATTGACA	AATGCTTAGT	TGAACCTCTT
401	TCATCAAGCG	ATGTTTCAGA	TAGACAAAAAA	CTTCAATGAT	TTGGATTTC
451	ACTTAAGGGA	AATCAATTAG	CAAAGACCCA	TTTATTAACA	GCTCTTC
501	CTCAAAAGCA	GTATTTCTTT	CAAGACGAAT	GGAACCAAGT	TAGAGCAATG
551	ATTGGAAATG	AGCTCTTCCG	ACATCTCTAC	ACTAAATATT	TAATATTCCA
601	GCGAACTTCT	GAAGGAACTC	TTGTTCAATT	TTGCGGGAAAT	AACGTTTTG
651	ATCATTTGAA	AGTCAACGAT	AAGTTTGACA	AAAAGCAAAA	AGGTGGAGCA
701	GCAGACATGA	ATGAACCTCG	ATGTTGATCA	ACCTGCAAAT	ACAATGTC
751	GAATGAGAAA	GATCACTTT	TCAACAAACAT	CAACGTGCCG	AATTGGAATA
801	ATATGAAATC	AAGAACCCAGA	ATATTTTATT	GCACTCATTI	TAATAGAAAT
851	AACCAATTCT	TCAAAAAGCA	TGAGTTTGTG	AGTAACAAAAA	ACAATATTC
901	AGCGATGGAC	AGAGCTCAGA	CGATATTAC	GAATATATT	AGATTAAATA
951	GAATTAGAAA	GAAGCTAAAAA	GATAAGGTTA	TCGAAAAAAT	TGCCTACATG
1001	CTTGAGAAG	TCAAAGATT	TAACCTCAAC	TACTATTTAA	CAAAATCTTG
1051	TCCTCTTCCA	GAAAATTGGC	GGGAACCGAA	ACAAAAAAATC	GAAAACCTTG
1101	TAAATAAAAC	TAGAGAAAGAA	AAGTCGAAGT	ACTATGAAGA	GCTGTTAGC
1151	TACACAAC	ATAATAATG	CGTCACACAA	TTTATTAATG	AATTTTCTA
1201	CAATATACTC	CCCAAAGACT	TTTGACTGG	AAGAAACCGT	AAGAATTTTC
1251	AAAAGAAAGT	TAAGAAATAT	GTGGAACCAA	ACAAGCATGA	ACTCATTAC
1301	AAAAACTTAT	TGCTTGAGAA	GATCAATACA	AGAGAAATAT	CATGGATGCA
1351	GGTTGAGACC	TCTGCAAAGC	ATTTTTATT	TTTGATCAC	GAAAACATCT
1401	ACGTCTTATG	GAAATTGCTC	CGATGGATAT	TCGAGGATCT	CGTCGTC
1451	CTGATTAGAT	GATTTTCTA	TGTCACCGAG	CAACAGAAAA	GTTACTCCAA
1501	AACCTATTAC	TACAGAAAAGA	ATATTTGGGA	CGTCATTATG	AAAATGTC
1551	TCGCAGACTT	AAAGAAGGAA	ACGCTTGCTG	AGGTCCAAGA	AAAAGAGGTT
1601	GAAGAAATGGA	AAAAGTCGCT	TGGATTTGCA	CCTGGAAAAC	TCAGACTAAT
1651	ACCGAAGAAA	ACTACTTTCC	GTCCAATTAT	GACTTTCAAT	AAGAAGATTG
1701	TAAATTCTAGA	CCGGAAGACT	ACAAAATTAA	CTACAAATAC	GAAGTTATTG
1751	AACTCTCACT	TAATGCTTAA	GACATTGAAG	AATAGAATGT	TTAAAGATCC
1801	TTTGGATTG	GCTGTTTTA	ACTATGATGA	TGTAATGAAA	AAGTATGAGG
1851	AGTTGTTG	CAAATGGAAG	CAAGTTGGAC	AACCAAAACT	CTTCTTGCA
1901	ACTATGGATA	TCGAAAAGTG	ATATGATAGT	GTAAACAGAG	AAAACATTC
1951	AACATTCTA	AAAACACTA	ATTACTTTC	TTCAGATTTC	TGGATTATGA
2001	CTGCACAAAT	TCTAAAGAGA	AAGAATAACA	TAGTTATCGA	TTGAAAAAC
2051	TTTAGAAAGA	AAGAAATGAA	AGATTATTTT	AGACAGAAAT	TCCAGAAGAT
2101	TGCACTTGA	GGAGGACAAAT	ATCCAACCTT	ATTCACTGTT	CTTGAAAATG
2151	ACAAAATGA	CTTAAATGCA	AAGAAAACAT	TAATTGTTGA	AGCAAAGCAA
2201	AGAAATTATT	TTAAGAAAGA	TAACCTTACTT	CAACCAGTCA	TTAATATTC
2251	CCAATATAAT	TACATTAAC	TTAATGGGAA	GTTTATAAA	CAAACAAAAG
2301	GAATTCCCTCA	AGGTCTTGA	GTTCATCAA	TTTGTCTAC	ATTTTATTAT
2351	GCAACATTAG	AGGAAAGCTC	CTTAGGATTC	CTTAGAGATG	AATCAATGAA

FIG. 9A



9/78

2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
 2451 TTTTGATTAC AACTCAAGAG ATAATGCAG TATTGTTTAT TGAGAAACTT
 2501 ATAAACGTAAGTCGTGAAAA TGGAATTTAAA TTCAATATGA AGAAACTACA
 2551 GACTAGTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
 2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
 2651 TCAATTGATA TGAAAACCTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
 2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
 2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTAATGAA TAACATTACC
 2801 CATTATTTA GAAAGACGAT TACAACCGAA GACTTGCAG ATAAAACCTCT
 2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
 2901 AATACAAGGA CCACTTTAACG AAGAACTTAG CTATGAGCAG TATGATCGAC
 2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
 3001 CCTTGTGTGC AATATTAAGG ATACAATTTC TGGAGAGGAG CATTATCCAG
 3051 ACTTTTTCT TAGCACACTG AAGCACTTA TTGAAATATT CAGCACAAAAA
 3101 AAGTACATT TCAACAGAGT TTGCATGATC CTCAAAGGCAA AAGAAGCAAA
 3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
 3201 CTATTCTAAC TTATTTGGAA AAGTTAATT TCAATTTTG TCTTATATAC
 3251 TGGGGTTTTG GGGTTTGGGG GTTTTGGGG

FIG. 9B

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD
 51 LEDIKIFQAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
 101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
 151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGA
 201 ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNNMKSRTTR IFYCTHFNRN
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAM
 301 LEVKVDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHELIH
 401 KNLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKLL RWIFEDLVVS
 451 LIRCFFYVTE QQKSYSKTY YRKNIWDVIM KMSIADLKKE TLAEVQEKEV
 501 EEWKKSLGFA PGKLRLIPKK TTFRPIMTFN KKIVNSDRKT TKLTNTKLL
 551 NSHMLMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN
 651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ
 701 RNYFKKDNLQ QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFY
 751 ATLEESSLGF LRDESMNPEN PNVNLLMRLL DDYLLITTQE NNAVLFIEKL
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
 851 SIDMKTLALM PNINLRIEGI LCTLNLMQT KKASMWLKKK LKSFLMNNIT
 901 HYFRKTITTE DFANKTLNKL FISGGYKYMQ CAKEYKDHFK KNLAMSSMID
 951 LEVSKIYIYV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK
 1001 KYIFNRVCMI LKAKEAKLKS DQCQSLIQYD A

FIG. 10



+

10/78

CCCCAAAACCCAAAACCCCTATAAGGGAGGGTAGTTAGA 60
 1 GGGGTTTGGGGTTTGGGGATATTTCTTTAACTCCATCAAATCT

a	P	Q	N	P	K	T	P	K	P	L	*	K	K	K	L	R	*	F	R	-	
b	P	K	T	P	K	P	Q	N	P	Y	K	K	R	K	N	*	G	S	L	E	-
c	P	K	P	Q	N	P	K	T	P	I	K	K	E	K	I	E	V	V	*	K	-

AATAAAATATTATTCCCGCACAAATGGAGATGGATATTGATTTGGATGATATAGAAAATT 120
 61 TTATTTATAATAAGGGCGTGTACCTACACTAAACCTACTATATCTTTAA

a	N	K	I	L	F	P	H	K	W	R	W	I	L	I	W	M	I	*	K	I	-
b	I	K	Y	Y	S	R	T	N	G	D	G	Y	*	F	G	*	Y	R	K	F	-
c	*	N	I	I	P	A	Q	M	E	M	D	I	D	L	D	D	I	E	N	L	-

TACTTCCTAATACATTCAACAAGTATAGCAGCTTGTAGTGACAAGAAAGGATGCAAAA 180
 121 ATGAAGGATTATGTAAGTTGTCATATCGTCGAGAACATCACTGTTCTTCCTACGTTT

a	Y	F	L	I	H	S	T	S	I	A	A	L	V	V	T	R	K	D	A	K	-
b	T	S	*	Y	I	Q	Q	V	*	Q	L	L	*	*	Q	E	R	M	Q	N	-
c	L	P	N	T	F	N	K	Y	S	S	S	C	S	D	K	K	G	C	K	T	-

CATTGAAATCTGGCTCGAAATCGCCTTCATTGACTATTCAAAGTTGCAAAAACAATTAG 240
 181 GTAACTTTAGACCGAGCTTAGCGGAAGTAACTGATAAGGTTCAACGTTTTGTTAATC

a	H	*	N	L	A	R	N	R	L	H	*	L	F	Q	S	C	K	N	N	*	-
b	I	E	I	W	L	E	I	A	F	I	D	Y	S	K	V	A	K	T	I	R	-
c	L	K	S	G	S	K	S	P	S	L	T	I	P	K	L	Q	K	Q	L	E	-

AGTTCTACTTCTCGGATGCAAATCTTATAACGATTCTTCTTGAGAAAATTAGTTAA 300
 241 TCAAGATGAAGAGCCTACGTTAGAAATATTGCTAAGAAAGAACTCTTTAATCAAATT

a	S	S	T	S	R	M	Q	I	F	I	T	I	L	S	*	E	N	*	F	*	-
b	V	L	L	L	G	C	K	S	L	*	R	F	F	L	E	K	I	S	F	K	-
c	F	Y	F	S	D	A	N	L	Y	N	D	S	F	L	R	K	L	V	L	K	-

AAAGCGGAGAGCAAAGAGTAGAAACATTACTAATGTTAAATAAAATCAGGTA 360
 301 TTTCGCCTCTCGTTCTCATCTTAACTTGTAAATGATTACAAATTATTTAGTCCATT

a	K	A	E	S	K	E	*	K	L	K	H	Y	*	C	L	N	K	I	R	*	-
b	K	R	R	A	K	S	R	N	*	N	I	T	N	V	*	I	K	S	G	N	-
c	S	G	E	Q	R	V	E	I	E	T	L	L	M	F	K	*	N	Q	V	M	-

TGAGGATTATTCTATTAGATCACTTCTTAAGGAGCATTATGGAGAAAATTACTAA 420
 361 ACTCCTAATAAGATAAAAAATCTAGTGAAGAATTCCCTCGTAATACCTCTTTAATGAATT

a	*	G	L	F	Y	F	L	D	H	F	L	R	S	I	M	E	K	I	T	*	-
b	E	D	Y	S	I	F	*	I	T	S	*	G	A	L	W	R	K	L	L	N	-
c	R	I	I	L	F	F	R	S	L	L	K	E	H	Y	G	E	N	Y	L	I	-

FIG. 12A

+



11/78

TACTAAAAGGTAAACAGTTGGATTATTCCTAGCCAACAATGATGAGTATATTAAATT
 421 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 480
 ATGATTTCCATTGTCAAACCTAATAAAGGGATCGGTTGTTACTACTCATATAATTAA

 a Y * K V N S L D Y F P S Q Q * * V Y * I -
 b T K R * T V W I I S L A N N D E Y I K F -
 c L K G K Q F G L F P * P T M M S I L N S -

 CATATGAGAATGAGTCAGGATCTGATACATCAGACTTACCAAGACAAACTCGCTAT
 481 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 540
 GTATACTCTTACTCAGTTCTAGAGCTATGTAGTCTGAATGGTTCTGTTGAGCGATA

 a H M R M S Q R I S I H Q T Y Q R Q T R Y -
 b I * E * V K G S R Y I R L T K D K L A I -
 c Y E N E S K D L D T S D L P K T N S L * -

 AAAACGCAAGAAAAAGTTGATAATCGAACAGCAGAAGAACTTATTGCATTTACTATTG
 541 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 600
 TTTTGCCTTCTTTCAAACATTAGCTGTCGCTCTGAATAACGTAATGATAAGC

 a K T Q E K V * * S N S R R T Y C I Y Y S -
 b K R K K K F D N R T A E E L I A F T I R -
 c N A R K S L I I E Q Q K N L L H L L F V -

 TATGGGTTTATTACAATTGTTTAGGTATCGACGGTGAACCTCCGAGTCTGAGACAAT
 601 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 660
 ATACCCAAAATAATGTTAACAAATCCATAGCTGCCACTTGAGGGCTCAGAACTCTGTTA

 a Y G F Y Y N C F R Y R R * T P E S * D N -
 b M G F I T I V L G I D G E L P S L E T I -
 c W V L L Q L F * V S T V N S R V L R Q L -

 TGAAAAAGCTGTTACAACGTGAAGGAATCGCAGTCTGAAAGTTCTGATGTATGCCAT
 661 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 720
 ACTTTTCGACAAATGTTGACTTCCTAGCGTCAAGACTTCAAGACTACACATACGGTA

 a * K S C L Q L K E S Q F * K F * C V C H -
 b E K A V Y N * R N R S S E S S D V Y A I -
 c K K L F T T E G I A V L K V L M C M P L -

 TATTTGTGAATTAATCTCAAATATCTTATCTCAATTAAATGGATAGCTATAGAAACAAA
 721 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 780
 ATAAAACACTTAATTAGAGTTATAGAATAGAGTTAAATTACCTATCGATATCTTGT

 a Y F V N * S Q I S Y L N L M D S Y R N K -
 b I L * I N L K Y L I S I * W I A I E T N -
 c F C E L I S N I L S Q F N G * L * K Q T -

 CCAAATAAACATGCAAGTTAATGAAATACGTTAACCTTGGGACAAATGCACAC
 781 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 840
 GGTTTATTGGTACGTTAACATTACCTATATGCAATTAGGAAACCTGTTACGTGT

 a P N K P C K F N G I Y V K S F G T N A H -
 b Q I N H A S L M E Y T L N P L G Q M H T -
 c K * T M Q V * W N I R * I L W D K C T L -

 TGAATTATATTGGATTCTTAAAGCATAGATAACACAGAATGCTTAGAGACTGATTAGC
 841 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
 ACTTAAATATAACCTAACGAAATTGCTATCTGCTTACGAAATCTGACTAAATCG

 a * I Y I G F L K H R Y T E C F R D * F S -
 b E F I L D S * S I D T Q N A L E T D L A -
 c N L Y W I L K A * I H R M L * R L I * L -

FIG. 12B

+



12/78

TTACAAACAGATTACCTGTTTGATTACTCTTGCATCTCTTATATCTTAAAAGAAGCA
 901 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 960
 AATGTTGTCTAATGGACAAAACATAATGAGAACGAGTAGAGAAATATAGAAATTCTCGT

a L Q Q I T C F D Y S C S S L I S L K E A -
 b Y N R L P V L I T L A H L L Y L * K K Q -
 c T T D Y L F * L L L L I S Y I F K R S R -

GGCGAAATGAAAAGAAGACTAAAGAAAGAGATTCAAAATTGTTGATTCTCTGTAACC
 961 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1020
 CCGCTTTACTTTCTTCTGATTCTTCTCTAAAGTTAAACAACTAAGAACAGACATTGG

a G E M K R R L K K E I S K F V D S S V T -
 b A K * K E D * R K R F Q N L L I L L * P -
 c R N E K K T K E R D F K I C * F F C N R -

GGAATTAAACAACAAGAATATTAGCAACGAAAAAGAAGAACAGACTATCACAAATCCTGATT
 1021 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080
 CCTTAATTGTTGTTCTTATAATCGTTGCTTTCTTCTCGATAGTGTAGGACTAAG

a G I N N K N I S N E K E E E L S Q S * F -
 b E L T T R I L A T K K K K S Y H N P D S -
 c N * Q Q E Y * Q R K R R R A I T I L I L -

TTAAAGATTTCAAAAATTCCAGGTAAAGAGAGATACATTCAATTAAATTATATTATAG
 1081 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1140
 AATTCTAAAGTTTAAGGCCATTCTCTATGTAAGTAATTAAAGTATATAATATC

a L K I S K I P G K R D T F I K I H I L * -
 b * R F Q K F Q V R E I H S L K F I Y Y S -
 c K D F K N S R * E R Y I H * N S Y I I V -

TTTTTCATTCACAGCTGTTATTTCTTTATCTTAACAATATTTTGATTAGCTGGAA
 1141 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200
 AAAAGTAAAGTGTGACAATAAAAGAAAATAGAATTGTTATAAAAAACTAATCGACCTT

a F F I S Q L L F S F I L T I F F D * L E -
 b F S F H S C Y F L L S * Q Y F L I S W K -
 c F H F T A V I F F Y L N N I F * L A G S -

GTAAAAAGTATCAAATAAGAGAACAGCGCTAGACTGAGGTAACCTAGCTTATTCACATTCA
 1201 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1260
 CATTTTCATAGTTATTCTCTCGCGATCTGACTCCATTGAATCGAATAAGTGTAGTA

a V K S I K * E K R * T E V T * L I H I H -
 b * K V S N K R S A R L R * L S L F T F I -
 c K K Y Q I R E A L D * G N L A Y S H S * -

AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTAAAAA
 1261 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1320
 TCTAGCTGGAAGTATAGGTTATGCTACTATTCCCTTGTGTCAGTAGGCAAAATT

a R S T F I Y P I R * * G N S S H P F * K -
 b D R P S Y I Q Y D D K E T A V I R F K N -
 c I D L H I S N T M I R K Q Q S S V L K I -

TAGTGCTATGAGGACTAAATTAGTCAAGAAATGGAGCCGAAATCTTAATCAAAA
 1321 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1380
 ATCACGATACTCCTGATTAAATCTCAGTTACCTCGGCTTAGAATTAGTTT

a * C Y E D * I F R V K K W S R N L N Q K -
 b S A M R T K F L E S R N G A E I L I K K -
 c V L * G L N F * S Q E M E P K S * S K R -

FIG. 12C

+



13/78

GAATTGCGTCGATATTGCAAAAGAACATCGAACTCTAAATCTTCGTTAATAAGTATTACCA
 1381 -----+-----+-----+-----+-----+-----+ 1440
 CTTAACGCGACTATAACGTTTCTTAGCTTGAGATTAGAAAGCAATTATTCATAATGGT

 a E L R R Y C K R I E L * I F R * * V L P -
 b N C V D I A K E S N S K S F V N K Y Y Q -
 c I A S I L Q K N R T L N L S L I S I T N -

 ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAAGAAATAAA
 1441 -----+-----+-----+-----+-----+-----+ 1500
 TAGAACTAACTAACTCTCTAACTGCTCCGTTGACGTGCTCTAGTAATTCTTTATTT

 a I L I D C R D * R G N C T E D H * R N K -
 b S * L I E E I D E A T A Q K I I K E I K -
 c L D * L K R L T R Q L H R R S L K K * S -

 GTAACCTTTATTAATTAGAGAATAAAACTAAATTACTAATATAGAGATCAGCGATCTCAA
 1501 -----+-----+-----+-----+-----+-----+ 1560
 CATTGAAAAATAATTAATCTCTTATTTGATTAAATGATTATCTAGTCGCTAGAAGTT

 a V T F I N * R I N * I T N I E I S D L Q -
 b * L L L I R E * T K L L I * R S A I F N -
 c N F Y * L E N K L N Y * Y R D Q R S S I -

 TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAAATACAAACCTGGTCAAAAT
 1561 -----+-----+-----+-----+-----+-----+ 1620
 AACTGCTTATTTGACTTGATTCAATCTGTTATTTTATGTTGGAACCAAGTTTA

 a L T K * K L N * S * T I K N T N L G Q N -
 b * R N K S * T K V R Q * K I Q T L V K I -
 c D E I K A E L K L D N K K Y K P W S K Y -

 ATTGAGGAAGGAAAAGAACCGAGTTAGCAAAAGAAAAATAAGGAATAAAATAAAATGA
 1621 -----+-----+-----+-----+-----+-----+ 1680
 TAACTCCTCCTTCTGGTCAATCGTTCTTTATTCCGTTATTATTTACT

 a I E E G K E D Q L A K E K I R Q * I K * -
 b L R K E K K T S * Q K K K * G N K * N E -
 c * G R K R R P V S K R K N K A I N K M S -

 GTACAGAAGTGAAGAATAAAAGATTATTTTTCAATAATTATTGAAAAGAGGGTT
 1681 -----+-----+-----+-----+-----+-----+ 1740
 CATGTCTTCACTTCTTATTTCTAAATAAAAAGTTATTAAATAACTTTCTCCCCAA

 a V Q K * R N K R F I F F N N L L K R G V -
 b Y R S E E I K D L F F S I I Y * K E G F -
 c T E V K K * K I Y F F Q * F I E K R G F -

 TTGGGGTTTGGGGTTTGGGG
 1741 -----+-----+-----+ 1762
 AACCCCAAAACCCCAAAACCC

 a L G F W G F G -
 b W G F G V L G -
 c G V L G F W -

FIG. 12D



14/78

2	EVDVQNQADNHGIHSALKTCEEIKEAKTLYSWI0KVI0CRNQS0SHYKDL	51
19	ELELEMQENQNDIQVRVK....IDDPKQY..LVNVTAACLLQEGSYQDK	62
52	EDIKIFQAQTNI0VATPRDYNEEDFKVIARKEVF.STGLMIELIDKCLVELL	100
63	DERRYIITKALL....EVAESDPEFICQLAVYIRNELYIRTTNYIVAF.	107
101	SSSDVSDRQKLQCFGQLKGNLAKTHLLTALSTQKQYFFQDEWNQVRAM	150
108CVVHKNTQPFI0KVNKA0LLP0N0LLEVCEFAQVLYI	144
151	IGNELFRHLYTKYLIFQRTSEGTLVQFCGNVFDHLKVNDKF0KKQKGGA	200
145	FDATEFKNLY.....LDRILSQDIRKELTRKCLQRCVRSKF	181
201	ADMNE...PRCCSTCKY0V0KNEKDHF0LN0N0V0PNWNNMKS0TRIFYCTHF	247
182	SEFNEYQLGKYCTES..QRKKT0MFRYLSVTNKQKWDQTKKK.....	220
248	NRNNQFFKKHEFVSNKNNISAMDRAQTIFTN0FRFNRI0K0L0K0VIEKI	297
221	.RKENLLT0K0QAI0KES0D0K0S0K0RETG.....DIMNVEDA0K0L0K0PAVMKKI	264
298	AYMLEKV0KDFNFNYYLTKSCPLPENWRERKQKIENLINKTREEKSKYYEE	347
265	AKRQNAMK.....KHMKA0KIP0NSTLES0KYLTFKD	294
348	LFSYTTDNKC0VTQF0INEFFYNILPKDFLTGRNRKNF0QKKV0KKY0VELNKHE	397
295	LIKFC0HISEP.....KERVY0KILGK0K0PKT0EEYKA0FGDSAS0APFN.PE	338
398	LIHKNLL0KINTREISWMQVETSAKHFY0FDHENIYVLWKL0R0WIFEDL	447
339	LAGKRMKIE0ISK0TW0N0ELSAKGNTAEVWDNL0ISSNQLPYMAMLRNLSN..	386
448	VVSLIR0C0FFY0VTEQQKSYSKTY0Y0R0K0N0I0D0V0IM0K0MS0IA0L0K0KETLAEVQE	497
387ILKAGVSD.....	394
498	KEVEEWKKSLGFAPGK0L0L0IPK0K0T0F0R0P0I0M0T0F0N0K0K0I0V0N0S0D0R0K0T0K0L0T0N0T0	547
395TTHS	398
548	KLLNSH0ML0K0T0L0K0N0R0M0F0K0D0P0F0G0F0A0V0F0N0Y0D0V0M0K0Y0E0E0F0V0C0K0W0Q0G0P0K0L0	597
399	IVINK.....ICEPKAVENSKM	415
598	FFATMDIEKCY0D0SVN0R0E0K0L0T0F0L0K0T0K0L0S0S0D0F0W0I0M0T0A0Q0I0L0K0R0K0N0I0V0D0	647
416	F..PLQFFSAIEAVN.EAVTKGFKAKK...REN0M0N0L0K0G0Q0I0E0A0V0K0E..V0V0	457
648	SKNFRK0KEM0K0D0Y0F0R0Q0K0F0Q0K0IA0E0G0Q0Y0P0T0L0F0S0V0L0E0N0E0Q0N0D0L0N0A0K0T0L0I0V0E	697
458	KTDEEK0K0D0M0.....ELEQTEE0G0F0V0K0V0N0E0G0I0G0K0Q0Y0I0N0S0I0E0L0A0K0	496
698	AKQRNYF0K0D0N0L0L0Q0P0V0I0C0Q0N0Y0I0F0N0G0F0K0Q0T0K0G0I0P0G0L0C0V0S0S0I0L0S0	747
497	IAVNKNL0D0E0I0K0G0H0T0A0I0F0S0D0V0G0S0M0S0T0M0S0G0G0A0K0Y0G0S0V0R0T0C0L0E0C0A0L0V0G0L0L0	546
748	FYYATLE0E0S0L0G0F0L0R0D0E0M0N0P0N0P0V0N0L0L0M0R0L0T0D0Y0L0I0T0Q0E0N0A0V0L0F0I	797
547	MVKQRCEK0S0F0Y0I0F0S0P0S0S0Q0C0N0K0C0Y0L0E0V0D0L0.....	576

FIG. 13A

+



15/78

798 EKLINVSRENGFKFNMKK.LQTSFPLSPSKFAKYGMDSVEEQNIVQDYCD 846
 | . | |
 577 PGDELRPSMQKLLQEKGKLGGG..TDFPYECIDEWTKNKTHVD 617
 | . | |
 847 WIGISIDMKTTLALMPNINLRIEGILCTLNLNMQTKKASMWLKKKLKSFLM 896
 | . | |
 618 NIVILSDMMIAEGYSDINVRGSSIVNSI.....KKYKDEVN 653
 | . | |
 897 NNITHYFRKTITTEDFANKTLNKLFISSGGYKYMCAKEYKD.HFKKNLAM 945
 | . |
 654 PNIKIF...AVDLEGYYG.....KCLNLGDEFNENNYYIKIFGM 687
 | . |
 946 SSMIDLEVSKIIYSVTRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHFIE 995
 | . |
 688 SDSI.....LKFISAKQGGA.....NMVE 706
 | . |
 996 IFSTKKYIFNRVC 1008
 | . |
 707 VI..KNFALQKIG 717

FIG. 13B

132 LSTQKQYFFQDEWNQVRAMIGNEL.FRHLYTGYLIFQRTSE..GTLVQFC 178
 :| ..| |
 1 MSRRNQ.....KKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQI 43
 | . |
 179 GNNVFDHLKVNDKFDDKKQKGGAADMNEPRCCSTCKYNVKNEKDHFLLNNIN 228
 :| . . | |
 44 KEEDLKLLKFKNQDQDGNSGNDDDEE.....NNSNKQQELLRRVN 84
 | . |
 229 VPKWNNMKSRTTRIFYCTHFNRNNQFFKKHEFVSNKNNISAMDRAQTIFTN 278
 :| . . | |
 85 QIKQQVQLIKK...VGSKVEKDLNLNEDENKKN 114
 | . |
 279 IFRFNRIRKKLDKVIEKIAYMLEVKDFNFNYLTKSCPLPENWRERKQ 328
 :| .
 115 GLSEQQVKEEQLRTITEEQVKYQNLVFNMHDYQLDLNESGGHRRRRTDY 164
 | . |
 329 KIENLINKTREEKSYYEELFSYTTDNKCUTQFINE.FFYNILPKDFLTG 377
 :| .
 165 DTEKWFEISHDQK.....NYVSIYANQKTSYCWWLKDYFNK 200
 | . |
 378 RNRKNFQKKVKKYVELNKHELIHKNLLKINTREISWMQVETSACKFYY 427
 :| .
 201 NNYDHLNVSINRLE..TEAEFYAFDDFSQTIKLTNNSSYQTVNID..... 242
 | . |
 428 FDHENIYVLWKLLRWI..FEDLVVSLIRCFYYVTEQQKSYSKTYYRKNI 475
 :| .
 243 VNFDNNLCILALLRFLLSLERFNILNIRSSY..TRNQYNFEKIGELLETI 290
 | . |
 476 WDViMKMSIAIDLKETLAEVQEKEVEEWKKSLGFAPGKLRLIPKKTTFRP 525
 :| .
 291 FAVVFSHR.....HLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQ 330
 | . |
 526 IMTFNKKIVNSDRKTTKLTNTKLLNSHMLKTLKNRMFKDPFGFAVFNY 575
 :| | . | . . . | . |
 331 VYSFSTDLKLD..TNKVQDYFKFLQEFPRLTHVSQQAIPVSATNAVENL 378

FIG. 14A

+



16/78

576	DDVMKKYEEFVCKWKQVGQPKLF.	FATMDIEKCYDS..VNREK	615
379	NVLLKKVKH ANNLNLVSIPTOFNFDYFVNQLQHLKLEFGLEPNILTQKQK		426
516	LSTFL. KTTKLLLSSDFWIMTAQILKRKNNI..VIDSKNFRKEMK		657
427	LENLLLSIKQSKNLKFLRLNFYTYVAQETSRKQILKQATTIKNLKNKNQ		476
558	DYFRQKFQKIALEGQYPTLFSVLEN..EQNDLNAKTLIVEAKQRNYFK		705
477	EETPETKDETPSESTSGMKFFDHLSELTELEDFSVN....LQATQEIY..		520
706	KDNLLQPVINICQYNYINFNGKFYKQTKGIPQGLCVSSILSSFYATLEE		755
521	.DSLHKLLIRSTNLKKFKLSYKEMEKSMDTFIDLKNI.....YETLNN		564
756	SSLGFLRDESMNPNPVNLLMRLTDDYLITQENNAVLFIEKLINVSR		305
565LKRCSVNISNPHGNISYELTN.....KDSTFYKFKLTLNQE		500
806	ENGFKFNMKKLQTSFPLSPSKFAKYGMDSVEEQNIVQDYCDWIGISIDMK		855
601	LQHAKYTFK..QNEFQFNNVKSAKIESSSLESLEDIDSCLCKSIASCKNLQ		648
856	TLALMPNINLRIEGILCTLNLNMQT..KKASMWLKK..KLKSFLMNNITH		901
649	NVNI.....IASLLYPNNIQKNPFPNKPNNLLFFKQFEQLKNLENVSINC		691
902	YFRKTI...TTEDFANKTLNKLFISSGGYKYMQCAKEYKDHFKKNLAMSSM		948
692	ILDQHILNSISEFLEKNKKIKAFILKRYYLQYYLDYTKLFKTLQQLPEL		741
949	IDLEVSKIISVT.....RAFFKYLVCNIKDT..IFGEEHY		982
742	NQVYINQQLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESSQTLQLIDF		791
983	PDFFLS TLKHFIEIFSTKKY IFNRVCMILKAKEAKLKSDQCQSLIQ		1028
792	DQNTVSDDSIIKKILESISESKYHHYRLRNPSQSSSLIKSENEEIQELLK		840

FIG. 14B

4	DIDLDDIENLLPNTFNKYSSSCSDKKGCKTLKSGSKSPSLTIPK.	47
617	NVKSAKIESSSLESLEDIDSCLCKSIASCKNLQNVNIIASLLYPNNIQKNP	666
48LQKQLEFYFSDANLYNDSFLRKVLKSGEQRVE....IETLLM	86
667	FNKPNNLLFFKQFEQLKNLENVSINCILDQHILNSISEFLEKNKKIKAFIL	716

FIG. 15



17/78

1	MEMDIDLDDIENL.....LPNTFNKYSSSCSDKGCKTLKSGSKSPS...	42
	: 	

FIG. 16

telomerase p43 LQK[REDACTED]LEFYFSDANLYND[REDACTED]RKLVLKSGEQRVEIETLLM
human La ICHQUEYYFGDFNLPRDKFLKEQI.KLDEGWVPLEIMIK
Xenopus LaA ICEQIEYYFGDHNLPRDKFLKQQI.LLDDGWVPLETMIK
Drosophila La ILROVEYYFGDANLNRDKFIREQIGKNEDGWVPLSVLVT
S. c. Lhp1p CLK[REDACTED]LEFYFSEFNFPYDRFLIRTTAEK.NDGWVPISTIAT

FIG. 18

1 aactcattta attactaatt taatcaacaa gattgataaa aagcagtaaa taaaacccaa
 61 tagatthaat ttagaaagta tcaattgaaa aatggaaatt gaaaacaact aagcacaata
 121 gccaaaagcc gaaaattgt ggtggact tgaatttagag atgcaagaaaa accaaaatga
 181 tatataatgtt agggtaaga ttgacgtcc taagcaat ctcgtgaacg tcactgcagc
 241 atggttttg taggaaggta gttactacta agataaagat gaaagaagat atatcatcac
 301 taaagcactt cttgaggtgg ctgagtcga tcctgagttc atctgctagt tggcagtcta
 361 catccgtaat gaactttaca tcagaactac cactaactac attgttagcat tttgttgtgt
 421 ccacaagaat actcaaccat tcatcgaaaa gtacttcaac aaagcagttac ttttgcttaa
 481 tgacttactg gaagtctgtg aatttgcata ggttctctat atttttgtat caactgaatt
 541 caaaaatttg tattttgtata ggtacttca ataagatatt cgtaaggaaac tcacttccg
 601 taagtgttta caaagatgcg tcagaagcaa gtttctgtaa ttcaacaaat actaacttgg
 661 taagtattgc actgaatccat aacgtaaagaa aacaatgttc cgttacccat cagttaccaaa
 721 caagtaaaag tgggattaaa ctaagaagaa gagaaaaagag aatctctaa ccaaacttta
 781 ggcaataaaag gaatctgaag ataagtccaa gagaaaaact ggagacataa tgaacgttga
 841 agatgcaatc aaggctttaa aaccagcagt tatgaagaaaa atagccaaga gatagaatgc
 901 catgaagaaa cacatgaagg cacctaaat tcctaactct accttggaaat caaagtactt
 961 gaccttcaag gatactacac actcttattgt gatcaacaag atttgtgagc ccaaggccgt
 1021 gatccttggg aaaaaatacc ctaagaccga agaggaatac aaagcagct ttgtgattc
 1081 tgcattctgca cccttcaatc ctgatggc tcggaaagcg atgaaggattt aatctctaa
 1141 aacatggaaat aatgaactca gtgaaaagg caacactgtc gaggtttggg ataattttaaat
 1201 ttcaagcaat taactccat atatggccat gttacgttaac ttgtcttaaca tcttaaaagc
 1261 cgggttttca gataactacac actcttattgt gatcaacaag atttgtgagc ccaaggccgt
 1321 tgagaactcc aagatgttcc cttttagtgc ttgttgcattt atgaaggatc ttaatgaagc
 1381 agttacttaag ggatcaagg ccaagaagag agaaaaatatg aatctttaaaat gtc当地atcga
 1441 agcgtttttttt gaaatgttcc aaaaaacccaa tgaagagaag aagatataatg agttggatgg
 1501 aaccgaagaa ggagaattttt taaaatgtttt ccaaggaaat gatggatccat acattaaactc
 1561 cattgaacctt gcaatcaaga tagcgtttttt caagaattttt gatggatccat aaggacacac
 1621 tgcaatcttc tctgtatgtt ctgtttttt gatggatccat atgtcaggatc gagccaaagaa
 1681 gtatggttcc gttcgtactt ctgtttttt gatggatccat ctggatgttga tggtaaaata
 1741 acgtttgttca aagtccatcat tttttttt gatggatccat agttctcaat gcaataatgt
 1801 ttacttagaa gttgatctcc ctggatgttga actccgtttt tctatgtttt aacttttgca
 1861 agagaaatggg aaacttgggtt gttgatgttga tttttttt gatggatccat atgaatggac
 1921 aatctttaaaat actcacatgtt gatggatccat tttttttt gatggatccat ttggcaggaaagg
 1981 atattcagat atcaatgttta gatggatccat tttttttt gatggatccat agcatcaaaa agtacaagga
 2041 tgaagtaat ccttaacatta gatggatccat tttttttt gatggatccat gaaaggatccat gaaagtgcct
 2101 taatcttagt gatgatgttca gatggatccat tttttttt gatggatccat atattcgatc tgagcgattc
 2161 aatctttaaaat ttcattttccat gatggatccat tttttttt gatggatccat ttatctttaaaa
 2221 ctttgcctt caaaaatgtt gatggatccat tttttttt gatggatccat taacaaaaat
 2281 ctcaccccccac tttttttt gatggatccat tttttttt gatggatccat ttatctttaaaaat
 2341 atttaagtttta ctttccatgtt gatggatccat tttttttt gatggatccat ttatctttaaaaat
 2401 aaagaacaaa aaagattttt a

FIG. 19



18/78

	Motif A	Motif B	Motif C	Motif D	Motif E
Consensus telomerase p123	h-hDh---h-h	h---+---OG---SP	h---Yhdhhh	Gh-h---K	h-hLgh-h
Dong (LINE)	GQPKLFFATMDIEKCYDSVNRERKLSTFLKTTKL-100-KFYKOTKGIFQQGLCVSSILSSFFYATLEESSLGL	KNRNLHCTYDDYKKAFDSIPHSMWLIQVLEYKIN-	-14-LMRLTDDYLITLITQENN-0-AVLFIEKLINVSRENGFKENMKKLQT-23-QDYCDWIGISI		
a1 S.c. (groupII)	FGGSNNWREVDLKKCFDTISHDLIKELKRYISD-	28-RQIAIKKGIVQGDSLSPLWFCLALNPLSHQLHNDR	-16-HLIYMDIKLYAKNDKE-0-MKKLIDTTTFSNDISMQFLGDKCKT-25-KCILYKYLGFQQ		
HIV-RT	LKKKSIVTVLDVGDAYFASVPLDEDFRKYTAFTIP-	26-HPVVGPRVCVQGAPTSPLACNVALLRDRRLAGLA	-55-YVRYADDLIGVLGSKN-2-KLIKRDLNFLNS.LGITLINEEKTLI- 4-ETPAREFLGYNI		
L8543.12	VPEELYFMKEDVKSCLYDSDIPRMECMRILKDALKN-	7-GIRYQYNVLFQGWKGSPAIQFOSSTMKILEPFRQN	- 4-IYQYMDLIVYGSHEIG-1-HRTKIEELRQHLLRWGLTTPDKKHOK- 0-EPPFLWMGYEL		
		68-KCYIREDGIFQGSSU.SAPIVDLVYDLDLLEFYSEFK	- 8-ILKIAADDFLTISTDQQQ.....VINIKRLAMGGFQRYNAKANR-41-IRSKSSKGIFR		

FIG. 17

+

+



19/78

MEIENNOAQOPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL
VNVTAACLLQEGSYYQDKDERRIITKALLEVAESDPEFICQLA
VYIRNELYIRTTTNYIVAFCVVHKNTQPFIEKYFNKAVLLPNDL
LEVCEFAQVLYIFDATEFKNLYLDRILSQDIRKELTFRKCLQRC
VRSKFSEFNEYQLGKYCTESQRKKTMFRLSVTNKQKWDQTKKK
RKENLLTKLQAIKESEDKSKEGTDIMNVEDAIAKALKPAVMKKI
AKRNQAMKKHMKAPIOPNSTLESKYLTFKDLIKFCHISEPKER
YKILGKKYKTEEEYKAAGDASAPFNPALAGKRMKIEISKTW
ENELSAKGNTAEVWDNLISSNQLPYMAMILRNLSNILKAGVSDTT
HSIVINKICEPKAVENSKMFPLQFFSAIEAVNEAVTKGFKAKKR
ENMNKLKGQIEAVKEVVEKTDEEKKDMELEQTEEGEFVKVNEGIG
KQYINSIELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGA
KKYGSVRTCLECALVLGLMVQRCEKSSFYIFSSPSSQCNKCYL
EVDPGDELPRPSMQKLLQEKGKLGGGTDFPYECIDEWTKNKTHV
DNIVILSDMMIAEGYSDINVRGSSIVNSIKKYKDEVNPNIKIFA
VDLEGYKCLNLGDEFNENNYYKIFGMSDSILKFISAKQGGANM
VEVIKNFALQKIGQK

FIG. 20

MSRRNQKKPQAPIGNETNLDVFLQNLEVYKSQIEHYKTQQQQIK
EEDLKLLKFKNQDDQDGNSGNDDDEENNSNKQQUELLRRVNQIKQ
QVQLIKKVGSKVEKDLNLNEDENKKNGLSEQVKEEQLRTITEE
QVKYQNLVFNMODYQLDLNESGHRRRRETDYDTEKWFEISHDQ
KNYVSIXANQKTSYCWLKDYFNKNYDHLNVSINRLETEAEFY
AFDDFSQTIKLTNNSYQTVNIDVNFDDNNLCILALLRFLLSLERF
NILNIRSSYTRNQYNEKIGELLETI FAVVFSHRHLQGIHLQVP
CEAFQYLVNSSSSQISVKDSQLQVYSFSTDLKLVDTNKVQDYFKF
LQEFPRLTHVSQQAIPVSATNAVENLNVLKKVHANLNVLVSIP
TQFNFDYFVNLQHQLKEFGLEPNILTQKQLENLLSIKQSKNL
KFLRLNFYTVAQETSRKQILQATTIKNLKNNKNQEETPETKD
ETPSESTSGMKFFDHLSELTELEDFSVNLQATQEIYDSLHKLLI
RSTNLKKFKLSYKYEIMEKSKMDTFIDLKNIYETLNNLKRCVNI
SNPHGNISYELTNKDSTFYKFKLTLNQELQHAKYTFKQNEFQFN
NVKSAKIESSSLESLEDIDSCKSIASCKNLQNVNIIASLLYPN
NIQKNPFPNKPNNLFKQFEQLKNLENVSINCILDQHILNSISEF
LEKNKKIKAFILKRYYLLQYYLDYTKLFKTLQQLPELNQVYINQ
QLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESQTLQLIDFD
QNTVSDDSIKKILESISESKYHYLRLNPSQSSLIKSENSEEIQ
ELLKACDEKGVLVKAYYKFPLCLPTGTYYDNSDRW

FIG. 22

MKILFEFIQDKLDIDLQTNSTYKENLKGHFNGLDEILTTCFAL
PNSRKIALPCPGDLSHKAVIDHCIIYLLTGELYNNVLTFGYKI
ARNEDVNNSLFCHSANVNVLKGAAWKMFHSLVGTYAFVDLLI
NYTVIQFNGQFFTQIVGNRCNEPHLPPKWWORSSSSATAAOIK
QLTEPVTKQFLHKLNNINSSSSFPYSKILPSSSSIKKLDLREA
IFPTNLVKIPQRLKVRINLTLQKLLKRHKRLNYVSILNSICPPL
EGTVLDLHSLSRQSPKERVLFKIIIVILQKLLPQEMFGSKKNKGK
IIKNLNLSSLLSPLNGYLPDFSLLKKLRKDFRWLFIISDIWFTH
NFENLNQLAICFISWLFRQLIPKIIQTFYCTEISSTVTIVYFR
HDTWNKLITPFFIVEYFKTYLVENNVCRNHSNTLSNFNHSKMRI
IPKKSNNEFRIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEY
LRNKRPTSFTKIYSPTQIADRIKEFKQRLKKFNNVLPELYFMK
FDVKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFNTNTGVL
KLFNVVNASRVPKPYELYIDNVRTVHLSNQDVINVVEMEIFKTA
LWVEDKCYIREDGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPS
QDTLILKLADDFLIISTDQQQVINIKKLAMGGFQKYNAKANRDK
ILAVSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMNFHIRSKS
SKGIFRSLIALFNTRISYKTIDTNLNSTNTVLMQIDHVVKNISE
CYKSAFKDLSINVTONMQFHSFLQRIIEMTVSGCPITKCDPLIE
YEVRFITLNGFLESLSNTSKFDNIILLRKEIQHLQAYIYIYI
HIVN

FIG. 23



20/78

1 tcaatactat taattaataa ataaaaaaaaa gcaaactaca aagaaaaatgt caaggcgtaa
 61 ctaaaaaaaag ccataggctc ctataggcaa tgaaacaaat cttgattttg tattacaaaa
 121 tctagaagtt tacaaaagcc agattgagca ttataagacc tagtagtaat agatcaaaga
 181 ggaggatctc aagctttaa agttcaaaaa ttaagattag gatggaaact ctggcaacga
 241 tcatgtatgat gaagaaaaaca actcaaataa ataataagaa ttattaagga gagtcaatta
 301 gattaatgtag caagttaat tgataaaaaa agttggttct aaggttagaga aagatttgaa
 361 tttgaacgaa gatggaaaca aaaagaatgg acyttctgaa tagcaagtga aagaagagta
 421 attaagaacg attactgaag aatagttaa gtattaaaat ttatgttta acatggacta
 481 ccagttagat ttaaatgaga gtgggtggcca tagaagacac agaagagaaa cagattatga
 541 tactgaaaaa tggttgaaa tatctcatga ccaaaaaaat tatgtatcaa ttacgc
 601 ctaaaagaca tcataattgtt ggtggcttaa agattatttt aataaaaaca attatgatca
 661 tcttaatgt aagcattaaaca gactggaaac tgaagccgaa ttctatgcct ttatgtgatt
 721 ttcacaaaaca atcaaaactta ctaataattc ttactagact gtaacatag acgttaatt
 781 tgataataat ctctgtatc tcgcattgtc tagattttta ttatctactaa aagattca
 841 tattttgaat ataagatctt cttatacaag aatattat aattttgaga aattttgtga
 901 gctacttgaa actatctcg cagttgtct ttctatcg cacttacaag gcattcatt
 961 acaagttcct tgcgaagcgt tctaataattt agttaactcc tcatacataaa ttacgtttaa
 1021 agatagctaa ttataggtat actctttctc tacagactta aaatttagttt acactaaca
 1081 agtccaaagat tattttaaat tcttataaga attccctcgat ttgactcatg taagcttaga
 1141 ggctatccca gttatgtcta ctaacgctgt agagaacccctt aatgttttac ttaaaaagg
 1201 caagcatgtc aatcttattt tagtttat ccttacatca ttcaattttt atttctactt
 1261 tggtaattta taacatttga aatttaggtt tgatttagaa ccaaatattt tgacaaaaca
 1321 aaagcttgaa aatctacttt tgagtataaa ataataaaaa aatctttaat tttaagatt
 1381 aaacttttac acctacgtt ctttataaaaa cttccagaaaaa cagatattaa aacaagctac
 1441 aacaatcaaa aatctcaaaa acaataaaaa tcaagaagaa actcctgaaa ctaaagatga
 1501 aactccaaagc gaaaggcacaat gttgtatgaa atttttttgcatcttctg aattaaccga
 1561 gcttgaagat tttagcgtt acttggtaagc taccctaaagaa attttatgata gcttgcacaa
 1621 acttttgatt agatcaaaa attttaaaaaa gttcaattttt aatctttaat atgaaatgg
 1681 aaagagataaa atggatacat tcataatgtc taagaatattt tttttttttt aactgacaaa
 1741 taaaagatgc tctgttataa tatcaaatcc tcataatggaa accttcttgc aatggatca
 1801 taaagattct actttttataa aatttaagct gacctttaaac taagaattat aacacgctaa
 1861 gtataacttt aagtagaaacg aatttttaattt taataacgtt aatggatca aatggatca
 1921 ttcccttcatc gaaagcttag aagatattga tagtcttgc aatcttattt cttcttgc
 1981 aaatttacaa aatgttataa ttatcgccag ttgtctatc cccaaataa tttagaaaaaa
 2041 tcctttcaat aagcccaatc ttcttattttt caagcaattt gaataattga aatggatca
 2101 aaatgtatct atcaactgtt ttcttgcattt gcatataattt aattttttt cagaatttct
 2161 agaaaaagaat aaaaaaataa aagcattcat ttggaaaaa tattttttt tacaatatta
 2221 tccttgattt actaaattat tttttttttt tttttttttt tttttttttt tttttttttt
 2281 cattaaatggaaat caatttggaaat aatttgcattt gttttttttt tttttttttt
 2341 ccacaagcaa aatgttttctt atgaaccattt atgtttttt atcaaaagaat catcctaaac
 2401 ccttttagctt atagattttt accaaaaacat tttttttttt tttttttttt tttttttttt
 2461 agaatctata tctgtatc aatgtttttt tttttttttt tttttttttt tttttttttt
 2521 cagtttaattt aatctgtttt aatgtttttt tttttttttt tttttttttt tttttttttt
 2581 aggtgttttta gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 2641 cgattacaat tcagatagat gttttttttt tttttttttt tttttttttt tttttttttt
 2701 tgaatatttc ttgtttttt tttttttttt tttttttttt tttttttttt tttttttttt
 2761 atatattttt gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt
 2821 aaaaaatcg

FIG. 21

+



21/78

Oxytricha
Eupletes

LCVSYILSSFYANLEENALQFLRKEESMDPEKPETNLLMRLT
LCVSSILSSFYATLEESSLGFLRDESMNPENPNVNLLMRLT

FIG. 24

ATTTATACTCATGAAAATCTTATTGAGTCATTCAAGACAAGCTTGACATTGATCTACA
GACCAACAGTACTTACAAAGAAAATTAAAATGTGGTCATTCAATGGCCTCGATGAAAT
TCTAACTACGTGTTCCGCACTACCAAATTCAAGAAAAATAGCATTACCATGCCTTCTGG
TGACTTAAGCCACAAAGCAGTCATTGATCACTGCATCATTAACCTGTTGACGGGCGAATT
ATACAACAAACGTACTAACATTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAAATAG
TCTTTTTGCCATTCTGCAAATGTTAACGTTACGTTACTGAAAGCGCTGCTTGGAAAAT
GTTCCACAGTTGGTCGGTACATACGCATTGTTGATTATTGATCAATTATAACAGTAAT
TCAATTAAATGGGCAGTTTCACTCAAATCGTGGGTAACAGATGTAACGAACCTCATCT
GCCGCCAAATGGGTCCAACGATCATCCTCATCATCCGCAACTGCTGCGBAAATCAAACAA
ACTTACAGAACCAAGTGACAATAACAATTCTACACAAGCTCAATATAAATTCCCTTC
TTTTTTCCATTAGCAAGATCCTCCTTCATCATCATCTATCATAAAAGCTAATGACTT
GAGAGAAGCTATTTCACAAATTGGTTAAAATCCTCAGAGACTAAAGGTACGAAT
TAATTGACGCTGCAAAGCTATTAAAGAGACATAAGCGTTGAATTACGTTCTATT
GAATAGTATTGCCCCACCATGGAAAGGGACCGTATTGACTTGTGCAATTGAGTAGGCA
ATCACCAAAGGAACGAGTCTGAAATTATCATTGTTATTACAGAAGTTATTACCCC
AGAAATGTTGGCTCAAAGAAAATAAGAAAATTATCAAGAATCTAAATCTTTATT
AAGTTACCCCTAAATGGCTATTACCATTTGATAGTTGTTACCAAGCACAATTGAAAAC
GGATTTCGGTGGTTGTTCAATTCTGATATTGGTTACCAAGCACAATTGAAAAC
GAATCAATTGGCGATTGTTCAATTCTGGCTATTAGACAACTAATTCCAAAATT
ACAGACTTTTACTGCACCGAAATATCTCTACAGTACAATTGTTACTTAGACA
TGACTTGGAAATAACCTTACACCCCTTTATCGTAGAATATTAAAGACGTACTTAGT
CGAAAACAACGTATGTAGAAACCATATAAGTTACACGTTGTCACATTCAATCATAGCAA
AATGAGGATTATACCAAAAAAGTAATAATGAGTCAGGATTATTGCCATCCCATGCAG
AGGGGCAGACGAAGAAGAATTACAATTATAAGGAGAATCACAAAATGCTATCCAGCC
CACTCAAAAATTAGAATACCTAACGAAACAAAAGGCCGACTAGTTTACTAAATATA
TTCTCCAACGCAAATAGCTGACCGTATCAAAGAATTAAAGCAGAGACTTTAAAGAAATT
TAATAATGTCTTACCAAGAGCTTATTGATGAAATTGATGTCAAATCTGCTATGATT
CATACCAAGGATGGAATGTATGAGGATACTCAAGGATGCGCTAAAAAATGAAAATGGTT
TTTCGTTAGATCTCAATTCTCAATACCAATTACAGGTGTATTGAAGTTATTAAATGT
TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTATACATAGATAATGTGAGGACGGT
TCATTATCAAATCAGGATGTTAAACGTTGAGAGATGGAATATTAAAACAGCTTT
GTGGGTGAAGATAAGTGTACATTAGAGAAGATGGCTTTTCAAGGCTCTAGTTTAC
TGCTCCGATCGTTGATTGGTGTATGACGATCTCTGGAGTTTATAGCGAGTTAAAGC
CAGTCCTAGCCAGGACACATTAAATTAAAACCTGGCTGACGATTTCCTTATAATCAC
AGACCAACAGCAAGTGTACATATCAAAAAGCTTGCATGGCGGATTCAAAATATAA
TGCAGAAAGCCAATAGAGACAAAATTAGCCGTAAGCTCCCAATCAGATGATGATACGGT
TATTCAATTGTCACATGACATATTGTTAAAGAATTGGAAGTTGAAACATTCAAG
CACAATGAATAATTCCATATCCGTTGAAATCTAGTAAAGGGATATTGCAAGTTAA
AGCGCTGTTAACACTAGAAATCTCTTATAAAACAATTGACACAAATTAAATTCAACAA
CACCGTCTCATGCAAATTGATCATGTTGAAAGAACATTCCGAATGTTATAATTCTGC
TTTAAGGATCTATCAATTAAATGTTACGCAAATATGCAATTTCATTGTTCTTACAACG
CATCATTGAAATGACAGTCAGCGGTTGTCACATTAGAAATGTGATCCTTAATCGAGTA
TGAGGTACGATTCAACCATATTGAATGGATTGGAAAGCCTATCTCAAACACATCAAA
ATTAAAGATAATATCATTCTTTGAGAAAGGAATTCAACACTGCAAGC

FIG. 26



22/78

human
tez1
EST2
p123

Motif 0
AKFLHWIIMSVVVVYELLRSFFYYVTTETTEQKNR
ISEIEWLVLGKRSNARMCLSDFEKRKQIFAEFIYWLNSFLIPIQSFFYVITESSDLRNR
LKDFRWLFISD--IWETKHNFENLNQLAICFISWLFRQLIPKIIQTFYYCTEISSTVT-
TREISWMQVET-SAKHFYYFDHEN-IYVLMKLLRWFEDLVVSLIRCFYYVTEQQKSYSK
* * *

human
tez1
EST2
p123

Motif 1
LFFYRKSWWSKLQSIGIROHLKRVQLRDVSEAVROHREARPALLTSRLRFIPKP--DGL
TVYFRKDIIWKLLCPEI-TSMKMEAFERKINENNVRMDTQK-TTLPPAVIRLLPKK--NTF
IVYFRHDWTWNKLITPPIFIVEYFKTYLVENNVCRNHNSYTLS--NFNHSKMRILIPKKSNNEF
YYYRKNIWDVIMKMSI-ADLKETLAEVQEKEVEEWKKS-LGFAPGKLRLIPKK--TTF
* * *

human
tez1
EST2
p123

Motif 2
RPIVNMDYVVGARTFRREKRAERLTSRVKALF-SVLYNERYA
RLITN-LRKRFLIKMGNSNKMLVSTNOTLRPVASILKHLINEESSGIPENLEVYMKLITF
RIIAIPCRGADEEEFTIYKENHNRNAIQOPTOKILEYLRNKRPTTSFTKIYSPTQIADRIKEF
RPIMTFNKKIVNSDRKTTKLTTNTKLINSHMLKTLKN-RMFKDPPGFAVFNYDDVMKRY
* * *

tez1
EST2
p123

Motif 3 (A)
KKDLIKHRMFGR-KKYFVRIDIKSCYDRIKQDLMFRIVKK-KLKDPFVIRKYATIHTATS
KQRLLKKFNVLPELYFMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFYRSQYFFNTN
EEFVCKWKQVGQPKLFFATMDIEKCYDSVNTREKLSTFLKTTKLSSDFWINTAQILKRN
* * *

FIG. 25

+

+



23/78

AKFLHWLMSVYVVELLRSFFYVTETTFQKNRLFFYRKSWSKLQSIGIRQHLKR
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVNMDYVGARTFREK
RAERLTSRVKALFSVLNYERA

FIG. 27

GCCAAGTTCTGCACTGGCTGATGAGTGTACGTCGAGCTGCTCAGGTC
TTCTTTATGTCACGGAGACCACGTTCAAAGAACAGGCCTTTCTACC
GGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAGACAGCACATTGAAG
AGGGTGAGCTCGGGACGTGTCGGAAGCAGAGGTCAAGCAGCATCGGAAGC
CAGGCCCGCCCTGCTGACTCCGCTTCATCCCCAAGCCTGACGGGC
TGCAGCCGATGTGAACATGGACTACGTCGTTGGAGCCAGAACGTTCCGCAGA
GAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTCAGCGTGCT
CAACTACGAGCGGGCGCG

FIG. 28

MTEHHTPKSRILRFLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLRSDVQTSFSIFLHSTVVGF
DSKPDEGVQFSSPKCSQSELIANVVQMFDESERRRNLLMKGFSMNHEDFRAMHNGVQNDLVSTF
PNYLISILESKNWLLEIIGSDAMHYLLSKGSIFEALPNDNLQISIPLFKNNVFEETVSKKRKR
TIETSITQNKSARKEVSWNSISISRFSIFYRSYKKFKQDLYFNLHSICDRNTVHMWLQWIFPRQFG
LINAFQVKQLHKVIPLVSQSTVVPKRLKVYPLIEQTAKRLHRISSLKVYNHCPYIDTHDDEKILS
YSLKPNQVFALRSILVRVFPKLIWGNQRIFEIILKDETLFLKLSRYESFLHYLMSNIKISEIEWL
VLGKRSNAKMCLSDFEKRQIFAEFIYWLNSIIIPILQSFFYITESSDLNRRTVYFRKDIWKLLCR
PFITSMKMEAFEKINENNVRMDTQKTLPPAVIRLLPKKNTFRLITNLRKRFLIKMGSNKMLVSTN
QTLRPVASILKHLINEESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDIKSCYDRIKQDLM
FRIVKKKLKDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPFEKVQVQLLSMKTSDTLFVDFVDY
WTKSSSEIFKMLKEHLSGHIVKIGNSQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL
LRVVDDFLFITVNKKDAKKFLNLSLRGFEKHNFSTSLEKTVINFENSGIINNTFFNESKRMPFFG
FSVNMRSLDTLLACPIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCN
IYRLGYSMCMRAQAYLKRMDIFIPQRMFITDLLNVIGRKIWKLAELGYSRRFLSSAEVKWLFC
LGMRDGLKPSFKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVLFLHRRiad

FIG. 29

+



24/78

FIG. 30A

-



25/78

TCTCTACTTTAAGAAAGGATCTTCTTAAGGCCACCGAATGTTGGGtaattataatgcgcgattccctcattattaattt
 gcaggGCTAAGAAGTATTGTAGGATAGATAATAAAATCCTGTTATGATCGAATAAGCAAGATTGATGTTTCGGATT
 GTTAaaaaAGAAACTCAAGGATCCCGAATTGTAAAGTATGCAACCATACATGCAACAAGTGACCAGCTACAAA
 AAACCTTGTGTTAGTGAGGCTTCCCTATGTTAGGAAATGTTAGGATTTCTCTTcatGGATAATTGTTAGTTGAT
 ATGGTGCCCTTGAAGAAGCATCTGCTGAGTTACTTTCTATGGATAACTGGATTTCTCTTcatGGATAATTGTTAGTTGAT
 GACCCAAAAGTTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAA
 attgttaataaactaaatgaaactagataAGGAaaATTCTCAAAAGTTGGTATCCCTCAGGGCTCAATTCTG
 CATTCTTTTGTCATTTCTATGGAAAGATTGATGTTAGGATAATCCATCGTTAACGTTAACGTTAACGTTAACGTT
 TTACGAGTAGTCGACGATTCTCTTATAACAGTTAACGTTAACAGTTAACAGTTAACAGTTAACAGTTAACAGTTAACAGG
 tggatgtcgatcattcataacgttcaacgttgaaggatttgaaaACACAAATTCTACAGGCCTGGAGAAAACAGTA
 ATAAACATTGAAAAATAGTAATGGGATAATAAAACATACTTTTTAATGAAAGCAAGAAAGATGCCATTCTGGTT
 CTCTGTGAACATGAGGTCTCTTGATACATTGTTAGCATGTCCTAAATTGATGAAAGCCTTATTAACCTACATCTGTAG
 AGCTGACGAAAACATAGGGAAATCTTCATAAATTCTAACGTTAACGTTAACGTTAACGTTAACGTTAACGTTAAC
 atcagatcggccTTGCATCTCTTGTGATACTCTATGTTGAGGGATACTTACCCACAAATTCAATTCAATTCTGCTGCAATAT
 ATATTAGGGCTAGGATACTCTATGTTGAGGGATACTTACCCACAAATTCAATTCAATTCTGCTGCTGCTGCAATAT
 TGTTICATAACGGGTggacttacttttaactagaaaagtcatraattaaaccttagatCTTTGAATGTTTATTGGAGAGAAA
 AATTGGAAAAAGTTGGCGAAATAATTAGGATAATAGCAATTAGGTTAGGGTTTCCTCTGCAAGTCAAATGGtaacgtgt
 cggctcgagacttcgaatattgacacatcggGTTTGTGAAATGAGAGATGGTTGAACACCCCTCTTTCCTTCAAA
 TATCATCCATGCTTCGACAGCTATAATACCAATTCTGATCTTACAGCTTATCAAGCCGCTAAGGACCAAGTTTGC
 ACAGGGCTTATTTTACATGAAAGAATAGCTGATTAAgtcattttcaattttcaattttcaattttcaattttcaatttt
 ttaacaacaatattactaagttatagttgacttt
 gttatt
 cccatttaaacggggatggtaaaccattaaacatggatggatggatggatggatggatggatggatggatggatggat
 tctataatgaaataatgcccggactaatgtggatggatggatggatggatggatggatggatggatggatggatggat
 aaaagagagtaatacccgatggatggatggatggatggatggatggatggatggatggatggatggatggatggat
 ttggtgaccggattttggtaaaaggccccaggatccatgtggccggccctgtctactggatggatggatggatggat
 agtttggatacttaatgcttataatgttt
 aaatagtgttaaggccattatgtggatccaaatagtggatccaaatgtggatccaaatgtggatccaaatgtggat
 aagctttagggcttcaaaaactccctgtatttaaaggaggaatcttccaccgtgaggaaatggatgttttttttttt
 gctggaggagaaggccctaatt
 ctccaggggatcccttgatgtcaataacttcttgcactgtcgcttcactgtcgactcttcgtactgtcgactct
 cgaggtaaggccaccaggatccatgtggatggatggatggatggatggatggatggatggatggatggatggatggat

F/G. 30B

A circular rubber stamp with a double-line border. The outer ring contains the text "U. S. PATENT & TRADEMARK OFFICE" at the top and "ST. LOUIS, MO." at the bottom. The inner circle contains the date "JUL 26 2004" in the center, surrounded by the serial number "5116325" which is curved along the inner edge.

26/78

FIG. 31

+



27/78

S-1: FFY VTE TTF QKN RLF FYR KSV WSK
S-2: RQH LKR VQL RDV SEA EVR QHR EA
S-3: ART FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q
A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS
A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

FIG. 32

Poly 4

5' - t a a g c c t c g
cag acc aaa gga att cca taa gg -3'
Q T K G I P Q G

4 (B')

5 (C')

3' - D D Y L L I T
ctg ctg atg gag gag tag tgg -5'
a a a a a a a a
t t t t
c c
Poly 1

FIG. 34

+

+



28/78

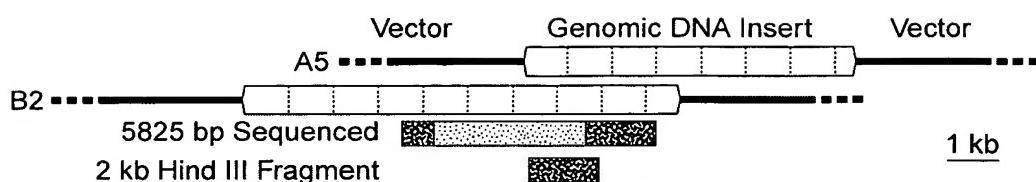


FIG. 33A

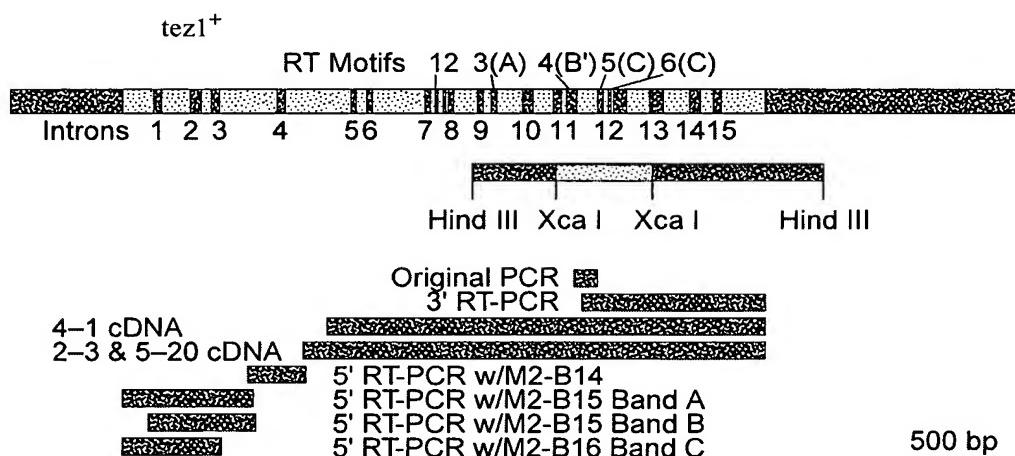


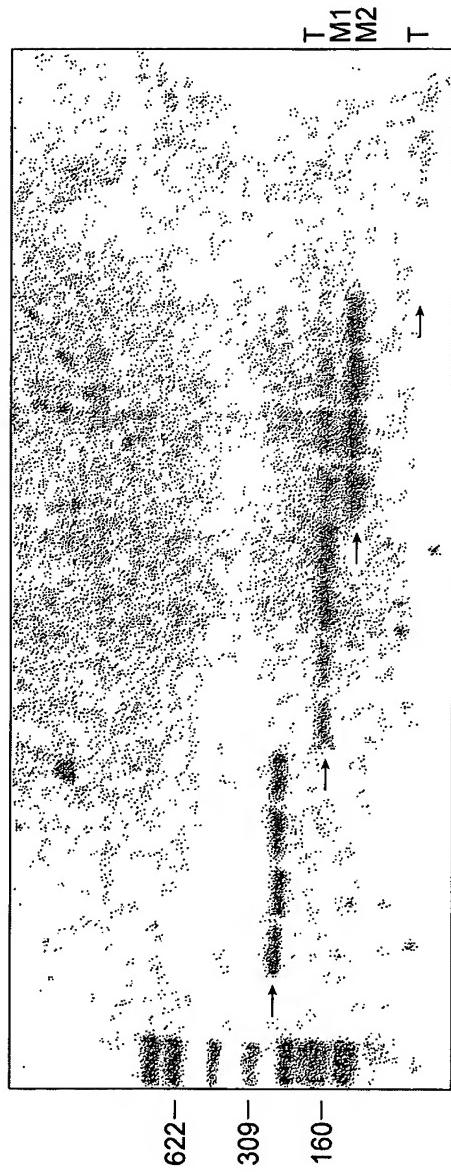
FIG. 33B

+

+



29/78



Motif B' (4)
QTKGIPQQ

Motif C (5)
DDYLLIT

FIG. 35

+



30/78

Ot	LCVSYIILSSFYANLEENALQFLRKESMDPEKPPETNLLMRLT
Ba_p123	KGIPQGLCVSSILSSFYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLIT
Sp_M2	SILSSFLCHFYMEDLIDEYLSTKKK-----GSVLLRVV
Sc_p103	DGLFQGSSLSSAPIVDLVYDDLLEFYSEFKASPS-----QDTLILKLAADDLIIIS
	* . * . *

Q K V G I P Q G
caa aaa gtt ggt atc cct cag gg..... <---Actual Genomic Sequence.

Poly 4
 t t t c
 t a a g c c t c g
 cag acc aaa gga att cca taa gg ----->

ag acc aaa gga att cca tca ggc tca ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATC
 tc tgg ttt cct taa ggt agt ccg AGT TAA GAC AGT AGA AAA AAC ACA GTA AAG ATA TAC
 K G I P S G S I L S S F L C H F Y M

F/G. 36A

+



31/78

GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC CCT AGT AAC AAT GCT
E D L I D E Y L S F T K K G S V L L R

GTA GTC gac tac ctc ctc atc acc
CAT CAG ctg ctg atg gag gag tag tgg

V V D D Y L L I T

<---- ctg ctg atg gag gag tag tgg
a a a a a a a a
t t t t t t
c c c c
Poly 1

.....gac gat ttc ctc ttt ata aca.....<---Actual Genomic Sequence
D D F L F I T

FIG. 36B

+

O P E R A T I O N S
JUL 26 2004
B I O T E C H N O L O G I C S
P O R T O R I C O

32/78

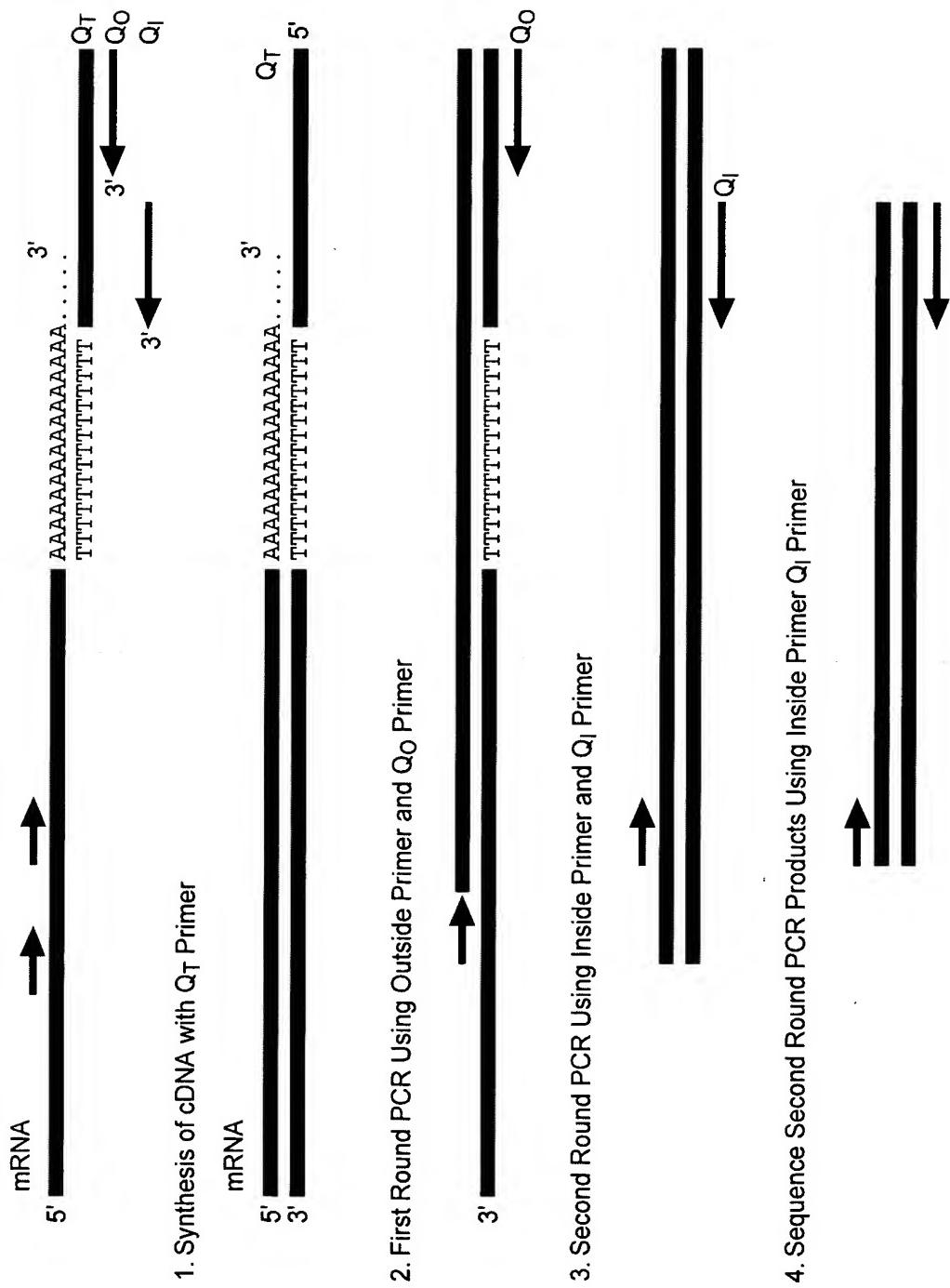


FIG. 37

+
JUL 26 2004
BUREAU OF INVESTIGATION & TRAILING MARK OFFICE 9115

33/78

- A. Genomic Libraries**
- Size Selected Libraries from P. Nurse
 - 3~4 kb
 - 5~7 kb
 - 7~8 kb
 - 11~12 kb
- B.**
- Libraries from J.A. Wise
- Sau 3a Partial Digest
 - Hind III Partial Digest
- cDNA Libraries
- GAD (Gal Activation Domain) Library
 - REP Library from R. Allshire
 - REP81ES Library (old)
 - REP81ES Library (new)
 - REP41ES Library

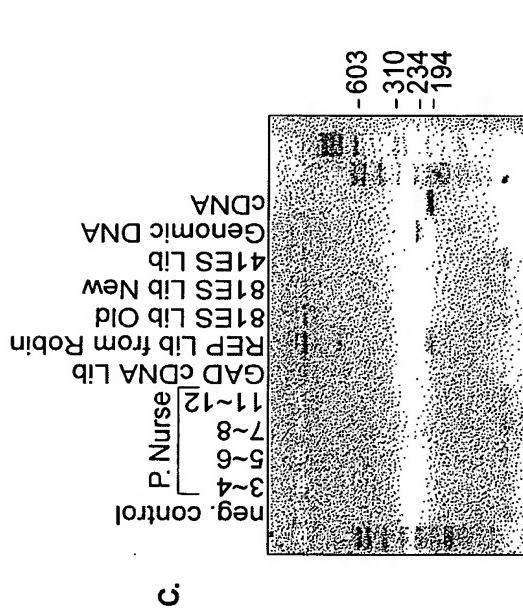
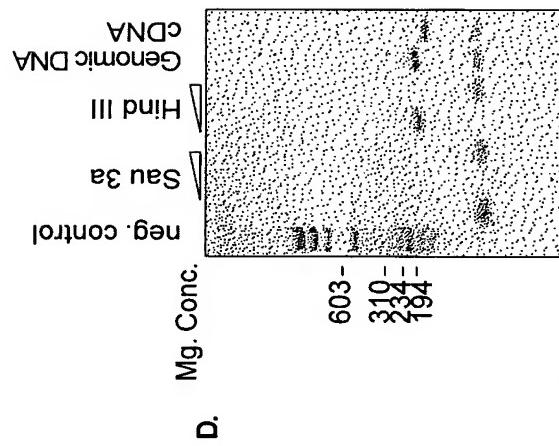
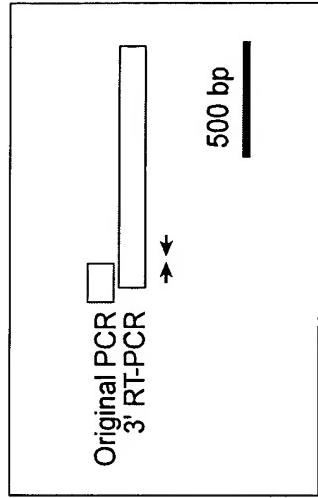


FIG. 38

+

O I P E JCT 16
JUL 26 2004
PATENT & TRADEMARK OFFICE

34/78

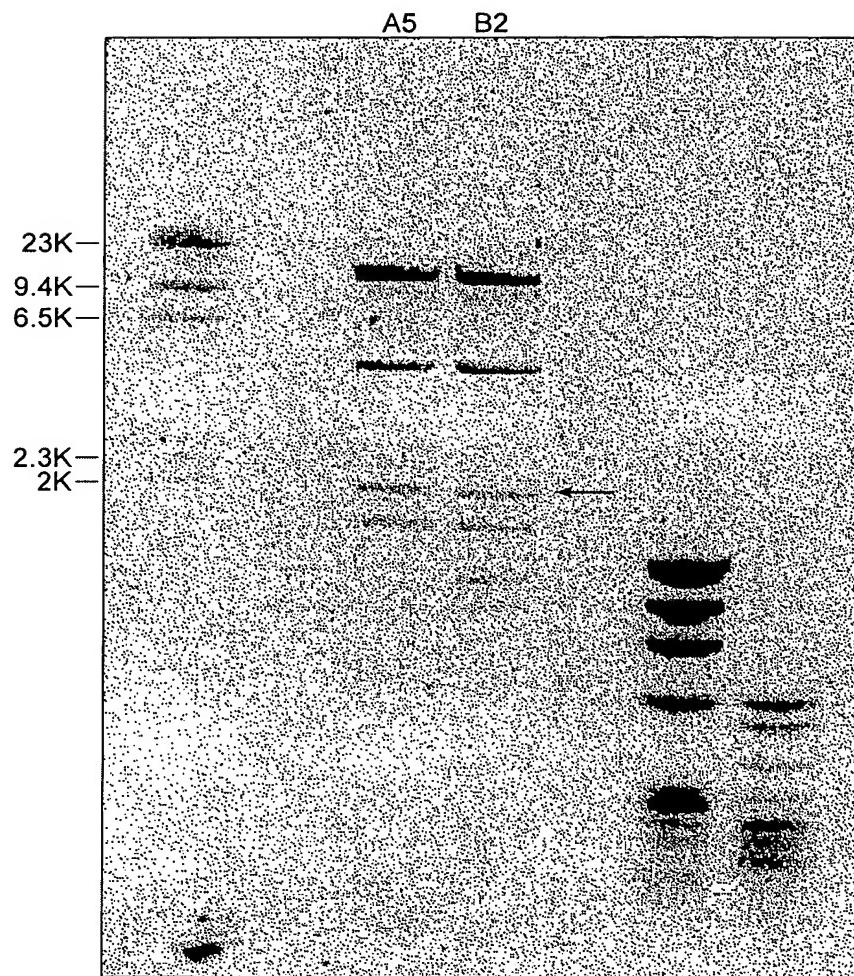
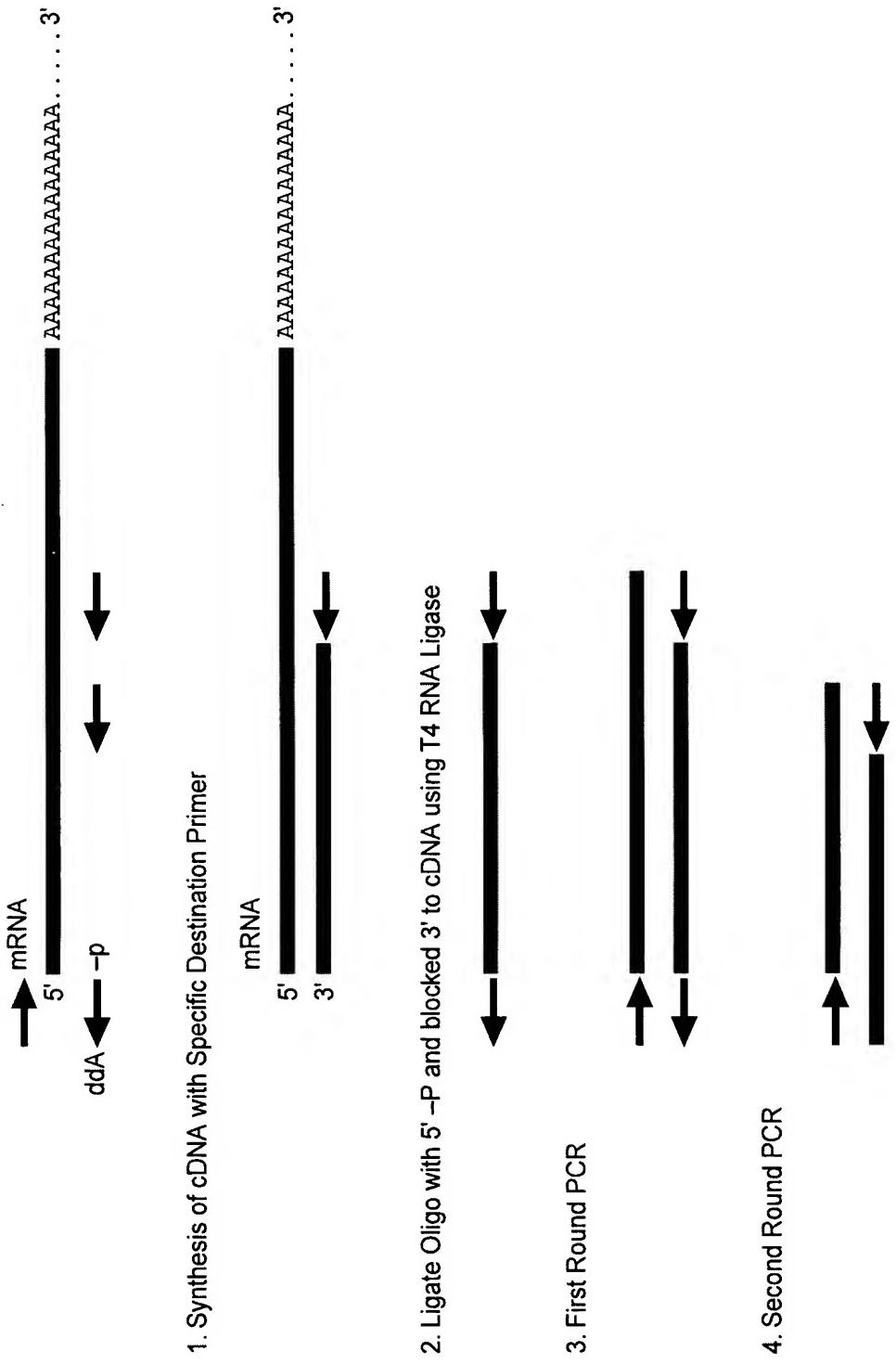


FIG. 39



35/78





36/78

		Motif 0					
S.p.	Tez1p	(429)	WLYNSFIIPILQSFYITESSSDLRNRTVYFRKDIW	...	(35)	...	
S.c.	Est2p	(366)	WLFRQLIPKLIQTFFYCTEISSTVT	- IVYFRHDTW	...	(35)	
E.a.	p123	(441)	WIFFEDLVVSILIRCFFYVTEQQKSYSKTYYRKNIW	...	(35)	...	
	*	*	*	*	*	*	*
		Motif 1			Motif 2		
		P	hh	K	K		
				hR	h	R	
S.p.	Tez1p	AVIRLLPKK--NTFRLITN-LRKRF	...	(61)	...		
S.c.	Est2p	SKMRIIPKKSNNNEFRIAIPCRGAD	...	(62)	...		
E.a.	p123	GKLRLIPKK--TTERPIMTFENKKIV	...	(61)	...		
	*	*	*	*	*	*	*
		Motif 3 (A) AF					
		h	hDh	GY	h		
S.p.	Tez1p	KKYFVRIDLICKSYDRIKQDLMFRIVK	...	(89)	...		
S.c.	Est2p	ELYFMKFDVKSCYDSIPRMECMRILK	...	(75)	...		
E.a.	p123	KLFFATMDIEKCYDSVNREKLSTFLK	...	(107)	...		
	*	*	*	*	*	*	*
		Motif 4 (B')					
		hPOG	PP	hh	h		
S.p.	Tez1p	YLQKVGTIPQGSILSSFLCHFYMEDLIDEYLSF	...	(6)	...		
S.c.	Est2p	YIREDGGLFQGSSLSSAPIVDLVYDDLFEFYSEF	...	(8)	...		
E.a.	p123	YKQTKGIGPOGLCVSSSILSSFYATTLEESSLGF	...	(14)	...		
	*	*	*	*	*	*	*
		Y Motif 5 (C)			Motif 6 (D)		
		h	F	DDhhh	Gh	cK	h
S.p.	Tez1p	VLLRVVDDFLFITVNKDAKKFLNLSLRGFEKHNFSTSLLEKTVINFENS	...	(205)	...		
S.c.	Est2p	LIIKLADDFFLIISTDQQVINIKKLAMGGFOQYNAKANRDKILAVSSQS	...	(173)	...		
E.a.	p123	LLMRLTDDYLITITQENNNAVLFIEKLINVSRENGFKFMKKLQTSEPLS	...	(209)	...		
	*	*	*	*	*	*	*

FIG. 41

+

A.

Sp_Tip1p 1 - - - - - MT EHHT PKSRI LRFLENQYVYLCT 24
 Sc_Est2p 1 - - - - - MEVDV DNQADNHGIH SALKTCEEIKEA KTLYSW 33
 Ea_p123 1

Sp_Tip1p 25 LND YVQL VLRGSSPA SYSN CERLRSDVQTSSFS 57
 Sc_Est2p 8 IQDK KLDIDLQTN - STYK - - ENLKCGHFNGLD 35
 Ea_p123 34 IQKV IRCRNQSQ - SHYK - - DLEDIKIEAQTN 61

Sp_Tip1p 58 IFLH STVVGFD SKPD EGVQFSSSPKC SQSEL IAN 90
 Sc_Est2p 36 EILT TCFALPNSR - KIAL PCLPGDL SHKAYIDH 67
 Ea_p123 62 IVAT PRDYN EEDFKV IARK EVFSTGLMIELIDK 94

Sp_Tip1p 91 VVKQMFD DESFERRR - NLLMKGF SMNHEDFRAMH 122
 Sc_Est2p 68 CIYLL TGELEYN - NVLT FGYKIARNED - - - 93
 Ea_p123 95 CLVELLS SSDVSDRQKLQCFGGFQLKGNNQ - - - 122

Sp_Tip1p 123 VNGVQNDL VSTFPNYLISI LESKNWQLLEIIG 155
 Sc_Est2p 94 - - VNNSL FCHSANVNVTLLKGAAWKMFHSSLVG 123
 Ea_p123 123 - - LAKTHLLTALSTQKQYFFQD EWNQVRAMING 152

Sp_Tip1p 156 SDAMHYLL SKGSIFEALPNNDNYLQISGIPLFKN 188
 Sc_Est2p 124 TYAFVDL LINYTQFN-GQFFTQIVGNRCNEP 155
 Ea_p123 153 NELFRHL YTQYLIFQRTSEGTLVQFCGNNVFDH 185

Sp_Tip1p 189 NVFEEETVSKK RKRTRITSITQN - - - KSARKEVS 218
 Sc_Est2p 156 HLPPPKWVQ - RSSSSATAAQI - - - KQLTEPV 183
 Ea_p123 186 LKVNDKF DK-KQKGGAADMNEPRCCSTCKNYVK 217

FIG. 42A

+



37/78

+

A.

Sp_Tip1p	219	WNSISISRFSSYK	KFQDLYFNLHSICD	251
Sc_Est2p	184	N-----NEK-	KQFLHKLNNSSSF	200
Ea_p123	218	DHFLLNNINVPNWNNMK	SRTTRIFYCTHFN	248

Sp_Tip1p	252	RNTVHMMWLNQWL	FPRQFFGL	I	NAFQVKQL	HKVIP	L	284
Sc_Est2p	201	-	-	-	-	-	SIKKLTDLREAIFP	223
Ea_p123	249	R	-	-	-	NNQFFKKHEFVSNKNNISAMDRAQTI		275

Sp_Tip1p	285	V S - - - Q S T V V P K R L L K V Y P L I E Q T A K R L H R I S	313
Sc_Est2p	224	T N - - - L V K I P Q R L K V R I N L T L Q K L L K R H K R L N	252
Ea_p123	276	F T N I F R E N R I R K K I K D K V I E F K I A Y M I F K V K D E N	308

Sp_Tip1p 314 L SKV Y NH Y CP Y ID - TH D D E K I L S Y S L K P N Q - - - 342
 Sc_Est2p 253 Y V S I L N S I C P P L E G T V L D L S H L S R Q S P K E R - - - 282
 Ea_r123 200 EN V Y I T K S C P P L E N W R E P K O K I E N I N K T R E K - - -

Sp_Tip1p 343 - - - - -
Sc_Est2p 283 - - - - -
Sc_Est2p 283 - - - - -
Sp_Tip1p 343 - - - - -

Sp_Tip1p 360 WGNQRIFEILKDLLETFLKL SRYESFSSLHYLMS 392
 Sc_Est2p 300 FGSKKNKGKIKNLLPLNGYL PFDSSLK 332

Sp_Tip1p	393	N I K I S E I E W L V L G K R S N A K M C L S D F E K R K Q I F A	425
Sc_Est2p	333	K L R L K D F R W L F I S - - D I W F T K H N F E N L N Q L A 362	
Ea_p123	407	K I N T R E I S W M Q V E T S - A K H E Y Y E D H E N - Y V W A 37	



38/78

+

+



39/78

A.

Sp_Tip1p	426	E F I Y W L Y N S F I I P I L Q S F F Y I T E S S D L R N R T V Y	458
Sc_Est2p	363	C F I S W L F R Q L I P K I Q T F F Y C T E I S S T V T - I V Y	394
Ea_p123	438	K L L R W I F E D L V V S L I R C F F Y V T E Q Q K S Y S K T Y Y	470
Sp_Tip1p	459	F R K D I W K L L C R P F I T S M K M E A F E K I N E N N V R M D	491
Sc_Est2p	395	F R H D T W N K L I T P F I V E Y F K T Y L V E N N V C R N H N S	427
Ea_p123	471	Y R K N I W D V I M K M S I A D L K K E T L A E V Q E K E V E E W	503
Sp_Tip1p	492	T Q K T T L P P A V I R L L P K K - N T F R L I T N L R K R F L	522
Sc_Est2p	428	Y T L S N F N H S K M R I I P K K S N N E F R I I A I P C R G A D	460
Ea_p123	504	K K S L G F A P P G K L R L I P K K - - T T F R P I M T F N K K I V	534
Sp_Tip1p	523	I K M G S N K K M L V S T N Q T L R P V A S I L K H L I N E - -	552
Sc_Est2p	461	E E E - - F T I Y K E N H K N A I Q P T Q K I L E Y L R N K R P T	491
Ea_p123	535	N S D - - R K T T K L T T N T K L L N S H L M L K T L K N R - M F	564
Sp_Tip1p	553	E S S G I P F N L E V Y M K L L T F K K D L L K H R M F G R - K K	584
Sc_Est2p	492	S F T K I Y S P T Q I A D R I K E F K Q R L L K K F N N V L P E L	524
Ea_p123	565	K D P F G F A V F N Y D D V M K K Y E E F V C K K W K Q V G Q P K L	597
Sp_Tip1p	585	Y F V R I D I K S C Y D R I K Q D L M F R I V K K K L K D P E - F	616
Sc_Est2p	525	Y F M K F D Y K S C Y D S I P R M E C M R I I K D A L K N E N G F	557
Ea_p123	598	F F A T M D I E K C Y D S V N R E K L S T F L K T K L L S S D F	630
Sp_Tip1p	617	V I R K Y A T I H A T S D R A T K N - - - - -	634
Sc_Est2p	558	F V R S Q Y F F N T N T G - - - - -	570
Ea_p123	631	W I M T A Q I L K R K N N I V I D S K N F R K K E M K D Y F R Q K	663

FIG. 42C

+

4

Sp_Tip1p	635	[F]V S E[A]F S Y F D M V P F E K V V Q L L S - - M K T S D T L F V	665
Sc_Est2p	571	- - - - - V L K L F N V V N A S R - - V P K P Y E L Y -	591
Ea_p123	664	[F]Q K I [A]L E G Q Y P T L F S V L E N E Q N D L N A K K T L I V	696
Sp_Tip1p	666	[D]F V D Y W T K [S] S S E - F K M L K E H L S G H I V K I G N S Q Y	698
Sc_Est2p	592	[D]N V R T V H L [S] N Q D V I N V V E M E I F K T A L W V E D K C Y	624
Ea_p123	697	E A K Q R N Y F K K D N L L Q P V I N I C Q Y N Y I N F N G K F Y	729
Sp_Tip1p	699	L [Q] K V G I P Q G S I L S S F L C H F Y M E D L I D E Y L S F T K	731
Sc_Est2p	625	I R E D G L F Q G S S I L S A P I V D L V Y D D L I E F Y S E F K A	657
Ea_p123	730	K Q T K G I P Q G L C V S S I L S S F Y Y A T L E E S S L G F L R	762
Sp_Tip1p	732	K K G - - - - - S V I L R V V D D F L F I T V N K K D A K K	756
Sc_Est2p	658	S P S Q D - - - - - T L I L K L A D D F L I S T D Q Q Q V I N	684
Ea_p123	763	D E S M N P E N P N V N I L L M R L T D D Y L L I T T Q E N N A V L	795
Sp_Tip1p	757	[E]L N L S L R G F E K H I N F S T S L E K T V I N F E N S N G - -	786
Sc_Est2p	685	I K K L A M G G F Q K Y N A K A N R D K I L A V S S Q S D - -	713
Ea_p123	796	[F]I E K L I N V S R E N G F K F N M K K L Q T S F P L S P S K F A	828
Sp_Tip1p	787	- - - I I N N T F F N E S K K R M P F F G F S V N M R S L D T L L	816
Sc_Est2p	714	- - - [D] D T V I Q F C A - - - M H I F V K E L E V W K H S S T M	739
Ea_p123	829	K Y G M D S V E E Q N I V Q D Y C D W I G I S I D M K T L A L M P	861
Sp_Tip1p	817	A C P K I [D] E A L F N S T S V E L T K H M G K S F F Y K I L R S S	849
Sc_Est2p	740	[N] N F H I R S K S S K G I F R S L I A L F N T R I S Y K T I D T N	772
Ea_p123	862	N I N L R I E G I L C T L N N M Q T K K A S M W L K K L K S F	894

40/78



FIG. 42D

+



41/78

A.

Sp_Tip1p	850	[L]SFAQV[FIDITHNS[KFN[SCCNIYR[LGYSMCMR 882
Sc_Est2p	773	[LNSTN[TYL[MQIDH[V[KN[SEC[- - - - - 793
Ea_p123	895	[LMNN[THYFRKTITTEDFANKTLNKLFISGGYK 927
Sp_Tip1p	883	AQAY[LKRMKDIFIPQRMFITDLLNVIGRKIWKK 915
Sc_Est2p	794	- - - YKSAFKDL[SIN-VTQNMQFHSFLQRIE 821
Ea_p123	928	YMQCAKEYKDH[EKKNLAMSSMIDLEVSKIIVSV 960
Sp_Tip1p	916	LAEILGYTSRRFLSSAEV[KWLFCLGMRDGGLKPS 948
Sc_Est2p	822	[TVSGCPIT[KCDPLIYEYEVRFTELNGFLESLSSN 854
Ea_p123	961	[TRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHF 993
Sp_Tip1p	949	FKYHPCFEQLIYQ[QSSLTD[LJ[KPLR[PVL[RQVLF 981
Sc_Est2p	855	TS- - - - - KFKDNILLRKEIQHLQAYIY 877
Ea_p123	994	IEIFS- - - TKK[YIFNRVCMLKAKEAKLKSDQC 1023
Sp_Tip1p	982	LHRR[AD- 988
Sc_Est2p	878	IYIHVN- 884
Ea_p123	1024	QSLIQYDA 1031

F/G. 42E

+

+

B.

Sp_Tip1p	1	- - - - -	M T E H H T P K S R I L R F L E N Q Y V Y L C T	24
Sc_Est2p	1	- - - - -	I Q D K L D I D L Q T N - - S T Y K - - E N L K C G H F N G L D	35
Ea_p123	1	M E V D V D N Q A D N H G I H S A L K T C E E I K E A K T L Y S W	33	

Sp_Tip1p	25	L N D Y V Q L V L R G S P A S S Y S N I C E R L R S D V Q T S F S	57
Sc_Est2p	8	I Q D K L D I D L Q T N - - S T Y K - - E N L K C G H F N G L D	35
Ea_p123	34	I Q K V I R C R N Q S Q - - S H Y K - - D L E D I K I F A Q T N	61

Sp_Tip1p	58	I F L H S T V V G F D S K P D E G V Q F S S P K C S Q Q S E L I A N	90
Sc_Est2p	36	E I L T T C F A L P N S R - K I A L P C L P G D L S H K A V I D H	67
Ea_p123	62	I V A T P R D Y N E E D F K V I A R K E V F S T G L M I E L I D K	94

Sp_Tip1p	91	V V K Q M F D E S F E R R R - N L L M K G F S M N H E D F R A M H	122
Sc_Est2p	68	C I Y L L T G E L Y N - - N V L T F G Y K I A R N E D - - -	93
Ea_p123	95	C L V E L L S S S D V S D R Q K L Q C F G F Q L K G N Q - - -	122

Sp_Tip1p	123	V N G V Q N D L V S T F P N Y L I S I L E S K N W Q L L E I I G	155
Sc_Est2p	94	- - V N N S L F C H S A N V N V T L L K G A A W K M F H S L V G	123
Ea_p123	123	- - L A K T H L L T A L S T Q K Q Y F F Q D E W N Q V R A M I G	152

Sp_Tip1p	156	S D A M H Y L L S K G S I F E A L P N D N Y L Q I S G I P L F K N	188
Sc_Est2p	124	T Y A F V D L L I N Y T V I Q F N - G Q F F T Q I V G N R C N E P	155
Ea_p123	153	N E L F R H L Y T K Y L I F Q R T S E G T L V Q F C G N N V F D H	185

Sp_Tip1p	189	N V F E E T V S K K R K R T I E T S I T Q N - - K S A R K E V S	218
Sc_Est2p	156	H L P P K W V Q - - R S S S S A T A A Q I - - K Q L T E P V T	183
Ea_p123	186	L K V N D K F D K - K Q K G G A A D M N E P R C C S T C K Y N V K	217

42/78



FIG. 42F

+

+



43/78

B.

Sp_Tip1p 219 **WNSISISRFSI** FYRSSYKKFKQDLYFNQLHSICD 251
Sc_Est2p 184 **N-----** -KQFLHKLNINSSFFP 200
Ea_p123 218 **NEK-----** DHFLNNINVPNWNMKSRTIFYCTHFN 248

Sp_Tip1p 252 **RNTVHMWLQWIFPRQFGLINAFAQVKQLHKVIP** L 284
Sc_Est2p 201 **TN-----** SIKKLTDLREAIFP 223
Ea_p123 249 **R-----** NNQFFKKHEFVSNKNNISAMDRAQT I 275

Sp_Tip1p 285 **VS-----** QSTVVVPKRLLKVVYPLIEQTAKRLHRI S 313
Sc_Est2p 224 **TN-----** LVKIPQRLLKVRINLTQKLLKRHKRLN 252
Ea_p123 276 **FTNIFRFNRIRKKLKDVKIEKIA** YMLEKVKDFN 308

Sp_Tip1p 314 **LSKVYNHYCPYID** THDDEKILSYSLKPNQ - - 342
Sc_Est2p 253 **YVSILNSICPPLEGTVLDLSHLSRQS** PKER - - 282
Ea_p123 309 **FNYYLTKSCP** LPENWRERKQKINELINKTREEK 341

Sp_Tip1p 343 **- - - - -** VFAFLRSILVRYFPKL I 359
Sc_Est2p 283 **- - - - -** VLKFIVILQKLLPQEM 299
Ea_p123 342 **SKYEEELFSYT** TDNKCYTQFINEFFYNI LPKDF 374

Sp_Tip1p 360 **WGNQRIFEI** ILKDLETFLKLSRYESFSLHYLMS 392
Sc_Est2p 300 **FGSKKNGK** IKNLNLLSPLNGYLPFDSSLK 332
Ea_p123 375 **LTG-RNRKNFQK** KVKKYVELNKHELIHKNLLE 406

Sp_Tip1p 393 **NIKISEI** EWLVLGKRSNAKMC LSDFEKRKQI FA 425
Sc_Est2p 333 **KLRLKDFRWLF** IS - - DIWFETKHNFENLNQLAI 362
Ea_p123 407 **KINTREI** SWMQVETS-AKHFYFYDHEN-IYVLW 437

FIG. 42G

+

+



44/78

B.

Sp_Tip1p	426	EF I YWL YNSFI PIL QSFFF I TE SSDL RNRT YY	468
Sc_Est2p	363	CF I SWL FRQL I PK I QT FF YCT E ISST VT - I VY	394
Ea_p123	438	KLLRW I FEDL VVSL I RCFF YVT EQQKS YSKT YY	470
Sp_Tip1p	459	FRKD I WKLL CRP F IT SMKMEA FEK IN ENN VRMD 491	
Sc_Est2p	395	FRHDTWNKL I TPFI VEYFKTYL VENN VCRNHNS 427	
Ea_p123	471	YRKN I WDV I MKMSIA DLKKET LAEVQEKEVEEW 503	
Sp_Tip1p	492	TQKTTLP PAV I RLL PKK - - NTFR L I TNLR KRF L 522	
Sc_Est2p	428	YTLSNFNHSKMR I PKK SNN EFR I IA I PCRGAD 460	
Ea_p123	504	KKSLGFAPGKLRL I PKK - - TTFRPIMTFNKKIV 534	
Sp_Tip1p	523	IKMGSNKKML VSTNQTLR PVAS I LKH LINE - - 552	
Sc_Est2p	461	EEE - - FTI YKENHHKNA I QPTQK I LEYLR NKR PT 491	
Ea_p123	535	NSD - - RKTTKLTTNTKLLN SHLMLKTLLKNR - MF 564	
Sp_Tip1p	553	ESSG I PFNL EY YMKLL TFKKD LL KHRMFGR - KK 584	
Sc_Est2p	492	SFTKI YSP TQI ADR I KEFKQRL L KKFNNVLPEL 524	
Ea_p123	565	KDPFGFA VFNYDDVMKKYEEF VCKWKQVGQQPKL 597	
Sp_Tip1p	585	YF VRI D I KSC YDR I KQDLMFRI VKKKLKDPE - F 616	
Sc_Est2p	525	YFMKFD YKSC YDSIPRMECMRI L KDALKNENGF 557	
Ea_p123	598	FFATMD I EKCYD SVNREKLSTFLK TTKLLSSDF 630	
Sp_Tip1p	617	VIRKYAT I HATSDRATKN - - - - - 634	
Sc_Est2p	558	FVRSSQYFFNTNTG - - - - - 570	
Ea_p123	631	WIMTAQI LKRKNN I VIDSKNFRK KEMKD YFRQK 663	

FIG. 42H

+

+

B.

Sp_Tip1p 635 FVSEAFSYFDMVPFEEKYVQLLS - MKTSDTLFV 665
Sc_Est2p 571 - - - VLKLFNVNNSR - VPKPYELYI 591
Ea_p123 664 FQKIALEGQQYPTLFSVLENEQNDLNAAKKTIV 696

Sp_Tip1p 666 DFVDYWTKSSSEIFKMLKEHLSGHIVKIGNSQY 698
Sc_Est2p 592 DNVRRTVHLSSNQDVINVVEMEIFKTAIWVEDKCY 624
Ea_p123 697 EAKQRNYFKKDNLQPVINICQYYINFNGKFY 729

Sp_Tip1p 699 LQKV[GIP[QGSILSSFLCHFYMEDLIDEYLSTDFK 731
Sc_Est2p 625 IREDGLFQGSSLSAPIVDLVYDDLEFFYSEFKA 657
Ea_p123 730 KQTKGIPQGLCVSSILSSFYATLEESSLGEELR 762

Sp_Tip1p 732 KKGG - - - SVLLRVVD[FLFIDTVNKDAKK 756
Sc_Est2p 658 SPSQD - - - TLLKLADDFLIISTDQQQVIN 684
Ea_p123 763 DESWNNPENPNVNLLMRLTDYLLITTQENNANVL 795

Sp_Tip1p 757 FLNLSLRGFEKHNFSTSLE[KTVINFENSNG - - 786
Sc_Est2p 685 IKKLAAMGGFQKYNAKANRDKILAVSSQSD - - 713
Ea_p123 796 FIEKLINVSRENGFKFNMKKLQTSFPLSPPSKFA 828

Sp_Tip1p 787 - - - INNTFFNESKKRMPFFGF SVNMRSSLDTLL 816
Sc_Est2p 714 - - - DDTVIQFCAMHIFVKELEVWVKHSSSTM 739
Ea_p123 829 KYGMDSVVEEQNI VQDYCDWIGISIDMKTLALMPP 861

Sp_Tip1p 817 ACPKIDEALFNSTSVELTKHMGKSFFFYKILRSS 849
Sc_Est2p 740 NNFHIRSKSSKGIFRSIALFNTRIISYKTIDTN 772
Ea_p123 862 NINLRIEGILCTLNLNMQTKKASMMWLKKLKSFP 894

FIG. 42I

+



45/78



46/78

B.

Sp_Tip1p	850	[LASFAQVFIDITHNSKFNSCCNIYRLGYSMCMR	882
Sc_Est2p	773	[LNSTNTVLMQIDHVVKNISEC	793
Ea_p123	895	[LMNNITHYFRKTITTEDFANKTLINKLFISGGYK	927
Sp_Tip1p	883	AQAYLKRWMKDIFIPQRMFITDLLNVIGRKIWKK	915
Sc_Est2p	794	- - YKSAFKDL S IN - VTQNMQFHSFLQRIEM	821
Ea_p123	928	YMQCAKEYKDHFKKNLAMSSMIDLEVS KIIYSV	960
Sp_Tip1p	916	LAEILGYT SRRFLSSAEVKWLFCLGMRDGLKPS	948
Sc_Est2p	822	TVSGCPITKCDPLIEYEVRFTILNGFLESLSSN	854
Ea_p123	961	TRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHF	993
Sp_Tip1p	949	FKYHPCFEQLIYQFQSLLTDLIKPLRPVLRQVLF	981
Sc_Est2p	855	TS - - - - KFKDNILLRKEIQHLQAYIY	877
Ea_p123	994	IEIFS - - TKKYIFNRVCMLKAKEAKLKSDQC	1023
Sp_Tip1p	982	LHRR I AD -	988
Sc_Est2p	878	IYIHIVN -	884
Ea_p123	1024	QSLIQYDA	1031

F/G. 42J

+

+

+



47/78

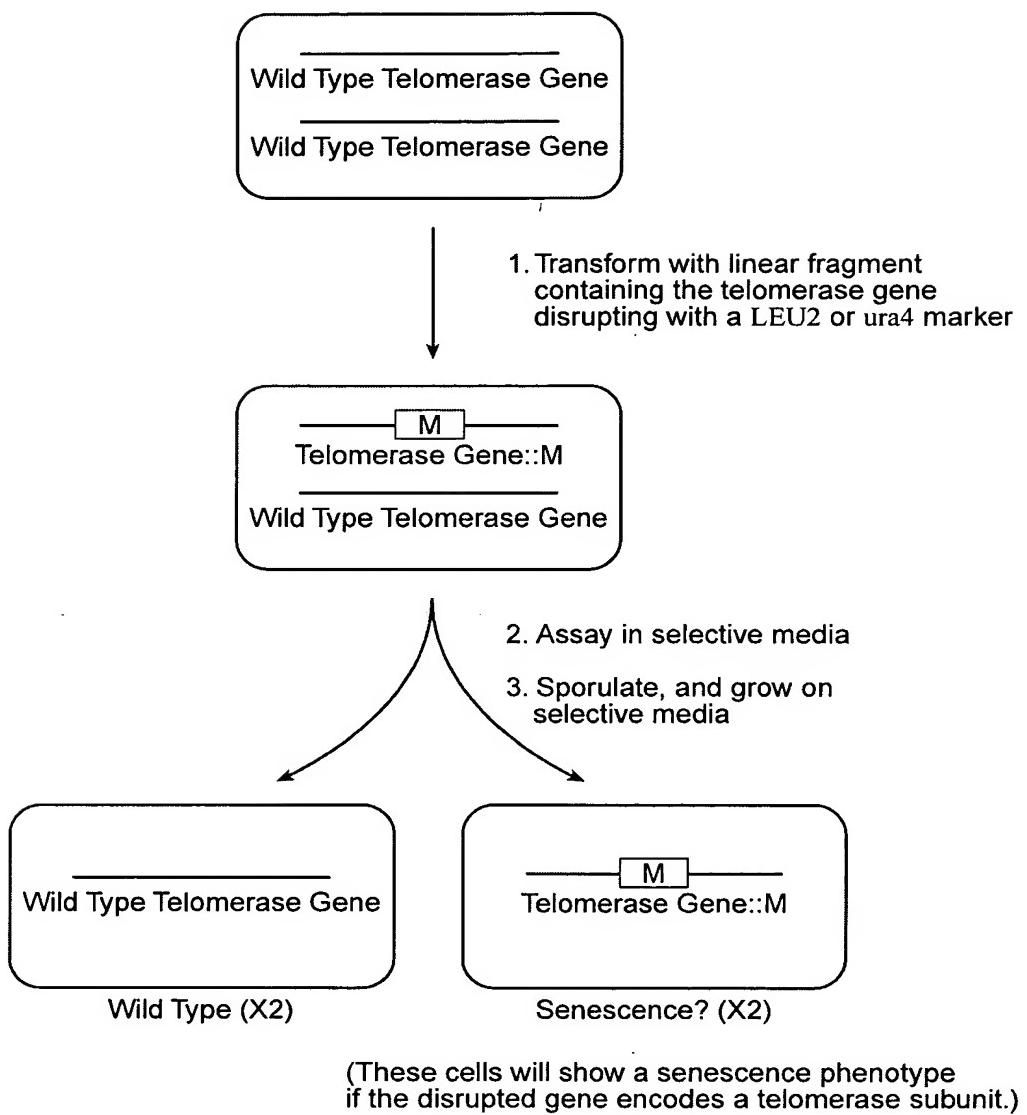


FIG. 43

+

+
BOSTON LIBRARIES
JULY 2004
48/78

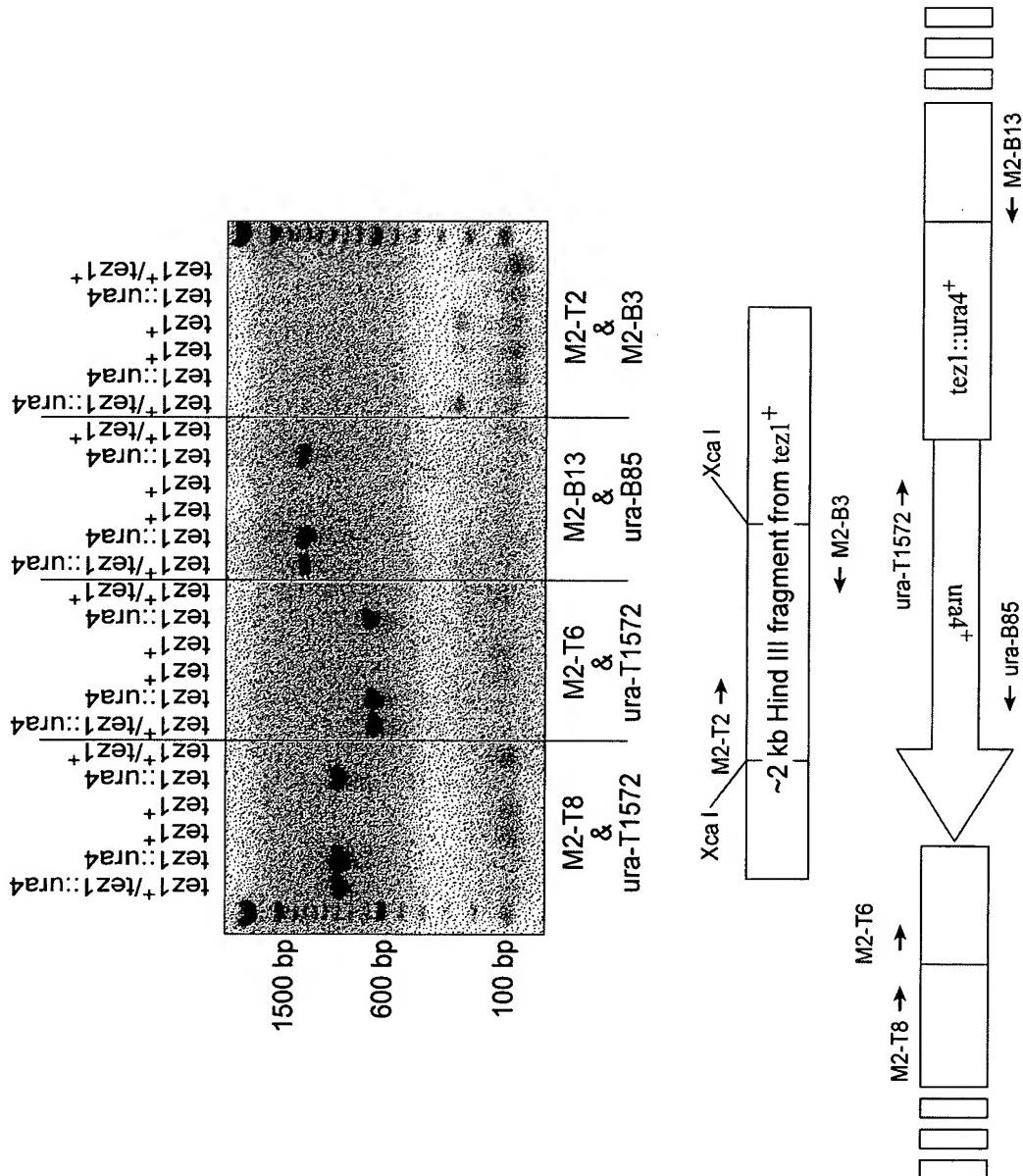


FIG. 44

O P E
JUL 26 2004
49/78
+
U.S. PATENT & TRADEMARK OFFICE
JC119

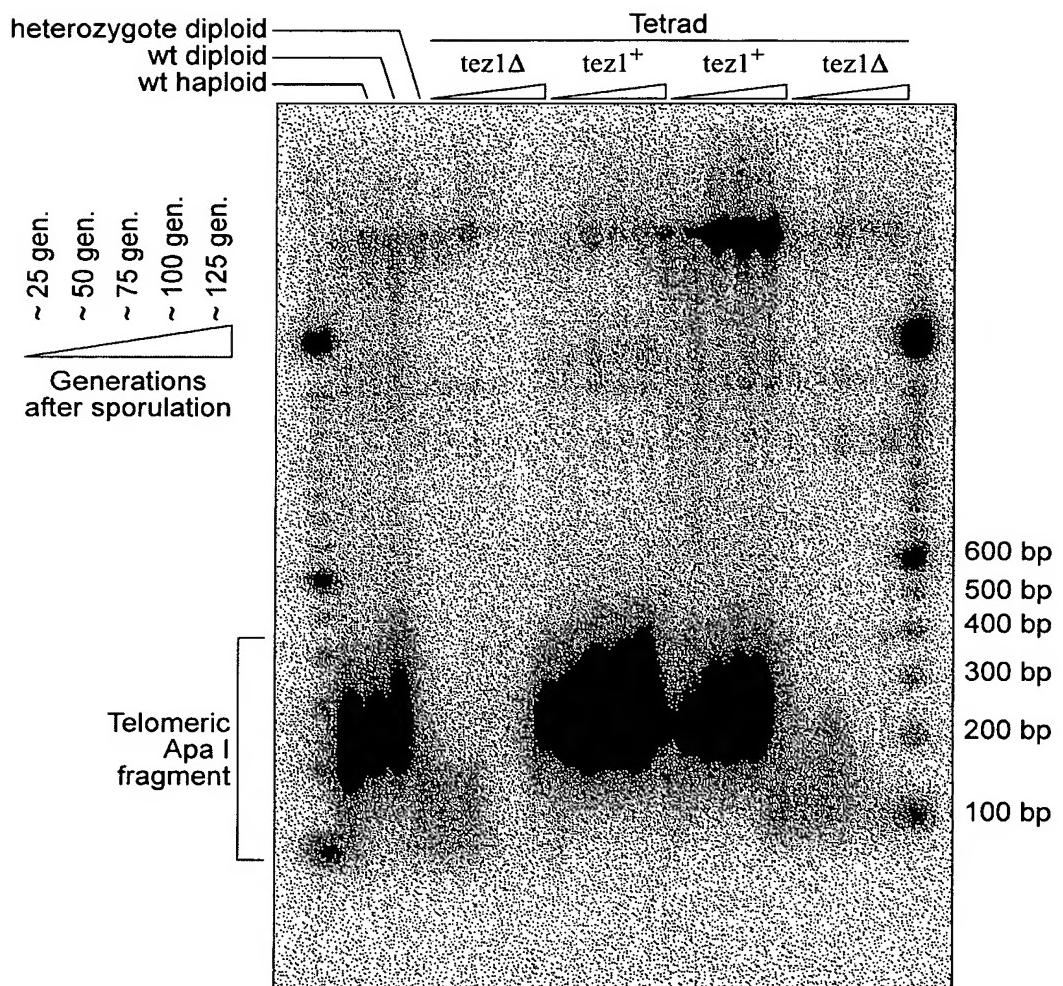


FIG. 45

+

FIG. 46A



50/78

-



51/78

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA 1529	
129 D L V S T F P N Y L I S I L E S K N W Q 148	
1530 CTT TTG TTA GAA AT gtaataccggtttaaggatgtgcacttgaacaagactgacaatag T ATC GGC 1601	
149 L L E I I G 155	
1602 AGT GAT GCC ATG CAT TAC TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661	
156 S D A M H Y L L S K G S I F E A L P N D 175	
1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721	
176 N Y L Q I S G I P L F K N N V F E E T V 195	
1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA 1781	
196 S K R K R T I E T S I T Q N T Q N K S A R K 215	
1782 GAA GTT TCC TGG AAT TCA ATT AGT AGG ATT TTT TAC AGG ATT TTT TAC AGG TCA TCC TAT 1841	
216 E V S W N S I S R F S I F Y R S S Y 235	
1842 AAG AAG TTT AAG CAA G gtaactaatactgttatccttcataactaaatttttag AT CTA TAT TTT AAC 1907	
236 K F K Q D L Y F N 245	
1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG ATT TTT CCA AGG 1967	
246 L H S I C D R N T V H M W L Q W I F P R 265	
1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA 2027	
266 Q F G L I N A F Q V K Q L H K V I P L V 285	
2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087	
286 S Q S T V V P K R L L K V Y P L I E Q T 305	
2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT AAC CAT TAT TGC CCA TAT ATT 2147	
306 A K R L H R I S L S K V Y N H Y C P Y I 325	
2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT ATT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207	
326 D T H D E K I L S Y S L K P N Q V F A 345	
2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267	
346 F L R S I L V R F P K L I W G N Q R I 365	

F/G. 46B

+



52/78

2268 TTT GAG ATA ATA TTA AAA G gtattgtataaaatccactaacgattttaccag AC CTC GAA ACT 2336
 366 F E I I L K D L E T 375

 2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396
 376 F L K S R Y E S F S L H Y L M S N I K 395

 2397 gtaatatgcacaaattttaccataacaatccag ATT TCA GAA ATT TCA GAA TGG CTA GTC CTT GGA 2465
 396 I S E I E W L V L G 405

 2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525
 406 K R S N A K M C L S D F E K R Q I F A 425

 2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585
 426 E F I Y W L Y N S F I I P I L Q S F F Y 445

 2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645
 446 I T E S S D L R N R T V Y F R K D I W K 465

 2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705
 466 L C R P F I T S M K M E A F E K I N E 485

 2706 gtattttaaagtatttttgcacaaaaggcttaataattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775
 486 N N V R M D T Q K T 495

 2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG ATT ACC TTT CGT CTC ATT ACG 2835
 496 T L P P A V I R L P K K N T F R L I T 515

 2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtattaaattttttggcatcaatgtactttacttcttaatcttata 2906
 516 N L R K R F I K

 2907 ttagccg ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967
 525 M G S N K K M L V S T N Q T L R P V 542

 2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027
 543 A S I L K H L I N E E S S G I P F N L E 562

 3028 GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT AAG CAC CGA ATG TTT GG gtaat 3088
 563 V Y M K L T F K D L K R M F G 581

FIG. 46C

+



53/78

3089	tataataatgcggattcatttttcgaa	G	CGT	AAG	AAG	TAT	TTT	GTA	CGG	ATA	GAT	ATA	3155								
582	R	K	K	Y	F	V	R	I	D	I			591								
3156	AAA	TCC	TGT	TAT	GAT	CGA	ATA	AAG	CAA	GAT	TTG	ATG	TTT	CGG	ATT	GTT	AAA	AAG	AAA	CTC	3215
592	K	S	C	Y	D	R	I	K	Q	D	L	M	F	R	I	V	K	K	K	L	611
3216	AAG	GAT	CCC	GAA	TTT	GTA	ATT	CGA	AAG	TAT	GCA	ACC	ATA	CAT	GCA	ACA	AGT	GAC	CGA	GCT	3275
612	K	D	P	E	F	V	I	R	K	Y	A	T	I	H	A	T	S	D	R	A	631
3276	ACA	AAA	AAC	TTT	GTT	AGT	GAG	GCG	TTT	TCC	TAT	T	gtaaggtttatttttcatggaaattttaacaa	3343							
632	T	K	N	F	V	S	E	A	F	S	Y	F									643
3344	attcttttag	TT	GAT	ATG	GTG	CCT	TTT	GAA	AAA	GTC	GTG	CAG	TTA	CTT	TCT	ATG	AAA	ACA	3405		
644	D	M	V	P	F	E	K	V	V	Q	L	L	S	M	K	T					659
3406	TCA	GAT	ACT	TTG	TTT	GTT	GAT	TTT	GTG	GAT	TAT	TGG	ACC	AAA	AGT	TCT	TCT	GAA	ATT	TTT	3465
660	S	D	T	L	F	V	D	F	V	D	Y	W	T	K	S	S	E	I	F		679
3466	AAA	ATG	CTC	AAG	GAA	CAT	CTC	TCT	GGA	CAC	ATT	GTT	AAG	gtataccatgttgaattttaacaa	3532						
680	K	M	L	K	E	H	L	S	G	H	I	V	K							692	
3533	ctaatgaaaactag	ATA	GCA	AAT	TCT	CAA	TAC	CTT	CAA	AAA	GTT	GGT	ATC	CCT	CAG	GGC	TCA	3593			
693	I	G	N	S	Q	Y	L	Q	K	V	G	I	P	Q	G	S				708	
3594	ATT	CTG	TCA	TCT	TTT	TIG	TGT	CAT	TTC	TAT	ATG	GAA	GAT	TTG	ATT	GAT	GAA	TAC	CTA	TCG	3653
709	I	L	S	F	L	C	H	F	Y	M	E	D	L	I	D	E	Y	L	S		728
3654	TTT	ACG	AAA	AAG	AAA	GGA	TCA	GTG	TTG	TTA	CGA	GTC	GAC	GAT	TTC	CTC	TTT	ATA	ACA	3713	
729	F	T	K	K	G	S	V	L	L	R	V	V	D	D	F	L	F	I	T		748
3714	GTT	AAT	AAA	AAG	GAT	GCA	AAA	AAA	TTT	TTG	AAT	TTA	TCT	TTA	AGA	G	gtgagtttgtgtcattcc	3777			
749	V	N	K	K	D	A	K	K	F	L	N	L	R	G						764	
3778	taagttctaacggtaag	GA	TTT	GAG	AAA	CAC	AAT	TTT	TCT	ACG	AGC	CTG	GAG	AAA	ACA	GTA	3840				
765	F	E	K	H	N	F	S	T	S	L	E	K	T	V						778	
3841	ATA	AAC	TTT	GAA	AAT	AGT	AAT	GGG	ATA	ATA	AAC	AAT	ACT	TTT	TTT	AAT	GAA	AGC	AAG	AAA	3900
779	I	N	F	E	N	S	N	G	I	I	N	N	T	F	N	E	S	K	K		798

F/G. 46D

+

54/78



FIG. 46E

+



55/78

FIG. 46F



56/78

1
met ser val tyr val val glu leu leu
GCCAAGTTCTGCACTGGCTG ATG AGT GTG TAC GTC GTC GAG CTG CTC

10
arg ser phe phe tyr val thr glu thr thr phe gln lys asn arg
AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG

20
leu phe phe tyr arg lys ser val trp ser lys leu gln ser ile
CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT

30
gly ile arg gln his leu lys arg val gln leu arg glu leu ser
GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG TCG

40
glu ala glu val arg gln his arg glu ala arg pro ala leu leu
GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG CTG

50
thr ser arg leu arg phe ile pro lys pro asp gly leu arg pro
ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG CCG

60
ile val asn met asp tyr val val gly ala arg thr phe arg arg
ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC AGA

70
glu lys ala glu arg leu thr ser arg val lys ala leu phe
GAA AAG ARG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG TTC

80
ser val leu asn tyr glu arg ala arg arg pro gly leu leu gly
AGC GTG CTC AAC TAC GAG CGG CGG CGC CCC GGC CTC CTG GGC

90
ala ser val leu gly leu asp asp ile his arg ala trp arg thr
GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC ACC

100
phe val leu arg val arg ala gln asp pro pro pro glu leu tyr
TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG TAC

110
phe val lys val asp val thr gly ala tyr asp thr ile pro gln
TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG

120
asp arg leu thr glu val ile ala ser ile ile lys pro gln asn
GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC AAA CCC CAG AAC

130
thr tyr cys val arg arg tyr ala val val gln lys ala ala met
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG

140
150
160
170
180
190
200

FIG. 47A



57/78

210

gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys
GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC

220

gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser
CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC

240

leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg
CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG

250

arg asp gly leu leu leu arg leu val asp asp phe leu leu val
CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG TTG GTG

270

thr pro his leu thr his ala lys thr phe leu arg thr leu val
ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC

280

arg gly val pro glu tyr gly cys val val asn leu arg lys thr
CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG AAG ACA

300

val val asn phe pro val glu asp glu ala leu gly gly thr ala
GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT

310

phe val gln met pro ala his gly leu phe pro trp cys gly leu
TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG

330

leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser
CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC

340

tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly
TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC

360

phe lys ala gly arg asn met arg arg lys leu phe gly val leu
TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG

370

arg leu lys cys his ser leu phe leu asp leu gln val asn ser
CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC

390

leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln
CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

400

ala tyr arg phe his ala cys val leu gln leu pro phe his gln
GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

FIG. 47B



58/78

420

gln val trp lys asn pro his phe ser cys ala ser ser leu thr
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430

440

arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

450

val ala gly gly gln gly arg arg arg pro ser ala leu arg gly
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460

470

arg ala val ala val pro pro ser ile pro ala gln ala asp ser
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

480

thr pro cys his leu arg ala thr pro gly val thr gln asp ser
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490

500

pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

510

pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520

530

his pro gly leu met ala thr arg pro gln pro gly arg glu gln
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

540

thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550

560

arg gly gly pro his pro gly leu his arg trp glu ser glu ala
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

564

OP

TGA GTGAGTGTGTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC
CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCTTCACTTCCCCAC
AGGCTGGCGTTGGTCCACCCCAGGGCCAGCTTCTCACCAGGAGGCCGGCTTCCACT
CCCCACATAGGAATAGTCATCCCCAGATTGCCATTGTTCACCCCTGCCCTGCCCTCC
TTTGCCTTCCACCCCCACCATTAGGTGGAGACCCCTGAGAAGGACCCCTGGAGCTTGGG
AATTGGAGTGACCAAAGGTGTGCCCTGTACACAGGGAGGACCCCTGCACCTGGATGGGG
GTCCCTGTGGGTCAAATTGGGGGAGGTGCTGTGGAGTAAAATACTGAATATATGAGTT
TTTCAGTTGGAAAAAAAAAAAAAA

FIG. 47C



59/78

Motif -1

Ep p123	...LVVSLIRCFYYVTEQQKSYSKT...
Sp Tez1	...FIIPILQSFFYITESSDLRNRT...
Sc Est2	...LIPKIIQTFFYCTEISSTVTIV...
Hs TCP1	...YVVELLRSFFYVTETTFQKNRL...
consensus	FFY TE

Motif 0

Ep p123	...KSLGFAPGKLRLIPKKT--TFRPIMTFNKKIV...	K R
Sp Tez1	...QKTTLPPAVIRLLPKKN--TFRLITNLRKRLF...	
Sc Est2	...TLSNFNHSKMRIIPKKSNNEFRIIAIPCRGAD...	
Hs TCP1	...ARPALLTSRLRFIPKPD--GLRPIVNMDYVVG...	
consensus	R PK R I	

Motif A

Ep p123	...PKLFFATMDIEKCYDSVNREKLSTFLK...
Sp Tez1	...RKKYFVRIDIJKSCYDRIKQDLMFRIVK...
Sc Est2	...PELYFMKFDVKSCYDSIPRMECMRILK...
Hs TCP1	...PELYFKVDVTGAYDTIPQDRLTEVIA...//...
consensus	F D YD

Motif B

Ep p123	...hPQG pS hh
Sp Tez1	...NGKFYKQTKGIPQGLCVSSILSSFYA...
Sc Est2	...GNSQYLQKVGIPQGSILSSFLCHFYME...
Hs TCP1	...EDKCYIREDGLFQGSSLSSAPIVDLVYD...
consensus	...RATSYVQCQGIPQGSILSTLLCSLCYG...
	G QG S

Motif C

Ep p123	...h F DD hhh
Sp Tez1	...PNVNLLMRLTDYLLITTQENN...
Sc Est2	...KKGSVLLRVVDDFLFITVNKKD...
Hs TCP1	...SQDTLILKLADDFLIISTDQQQ...
consensus	...RRDGLLLRLVDDFLLVTPHLTH...
	DD L

Motif D

Ep p123	...Gh h cK
Sp Tez1	...NVSRENGFKFNMKKL...
Sc Est2	...LNLSLRGFEKHNFST...
Hs TCP1	...KKLAMGGFQKYNAKA...
consensus	...LRTLVRGVPEYGCVV...
	G

FIG. 48

O I P E
JUL 26 2004
BOSTON OFFICE 9113

60/78

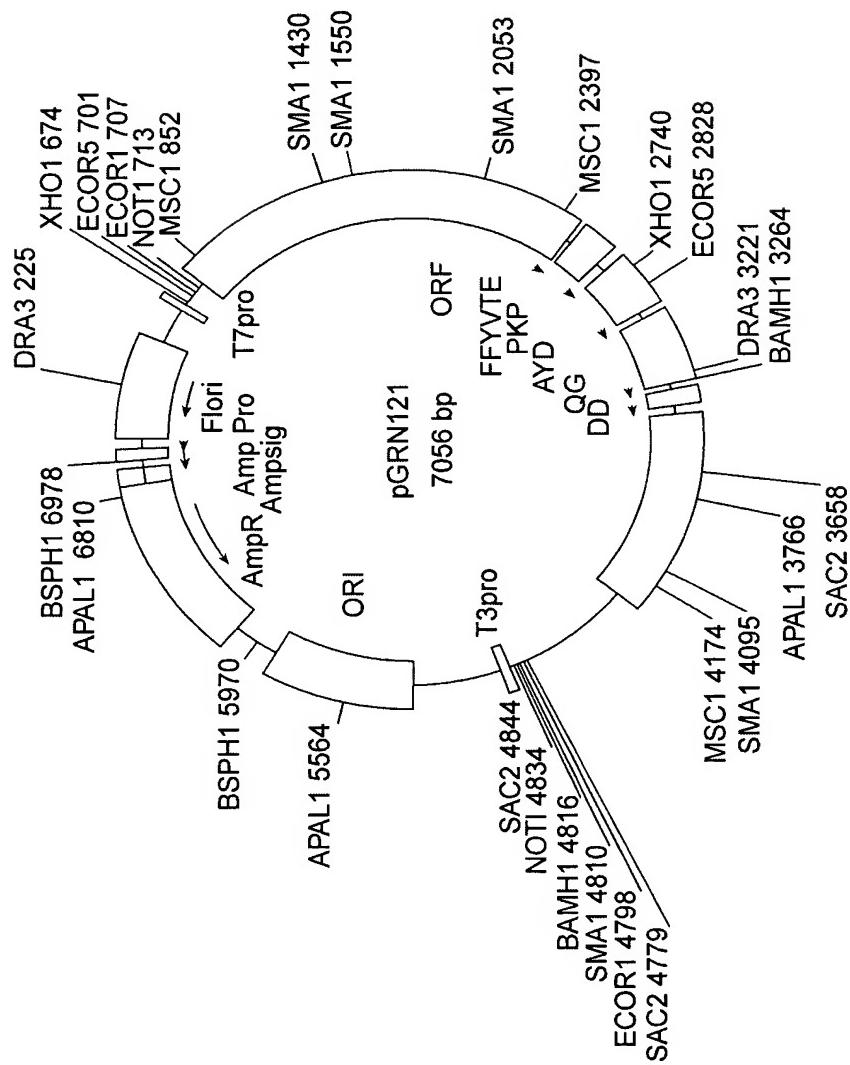


FIG. 49

+



61/78

1	GCAGCGCTGC	GTCCTGCTGC	GCACGTGGGA	AGCCCTGGCC	CCGGCCACCC
51	CCCGGATGCC	GCGCGTCCC	CGCTGCCGAG	CCGTGCGCTC	CCTGCTGCGC
101	AGCCACTACC	GCGAGGTGCT	GCCGCTGGCC	ACGTTCGTGC	GGCGCCTGGG
151	GCCCCAGGGC	TGGCGGCTGG	TGCAGCGCGG	GGACCCGGCG	GCTTCCGCG
201	CGNTGGTGGC	CCANTGCNTG	GTGTGCGTGC	CCTGGGANGN	ANGGCNGCCC
251	CCCGCCGCC	CCTCCTCCG	CCAGGTGTCC	TGCCTGAANG	ANCTGGTGGC
301	CCGAGTGCTG	CANANGCTGT	GCGANCACGG	CGCGAANAAC	GTGCTGGCCT
351	TCGGCTTCGC	GCTGCTGGAC	GGGGCCCGCG	GGGGCCCCCC	CGAGGCCTTC
401	ACCAACCAGCG	TGCGCAGCTA	CTGTGCCCAC	ACGGTGACCG	ACGCACGTGCG
451	GGGAGCGGGG	GCGTGGGGGC	TGCTGCTGCG	CCGCGTGGGC	GACGACGTGC
501	TGGTTCACCT	GCTGGCACCG	TGCGCGNTNT	TTGTGCTGGT	GGNTCCCGAC
551	TGCGCCTAAC	ANGTGTGCGG	GGCGCCGCTG	TACAGCTCG	GCGCTGCNAC
601	TCAGGGCCGG	CCCCGCCAC	ACGCTANTGG	ACCCGAANGC	GTCTGGGATC
651	CAACGGGCCT	GGAACCATAG	CGTCAGGGAG	GCCGGGGTCC	CCCTGGGCTG
701	CCAGCCCCGG	GTGCGAGGAG	GCGCGGGGGC	AGTGCAGGC	GAAGTCTGCC
751	GTTGCCAAG	AGGCCAGGC	GTGGCGCTGC	CCCTGAGCCG	GAGCGGACGC
801	CCGTTGGGCA	GGGGTCTCTGG	GCCCACCCGG	GCAGGACGCC	TGGACCGAGT
851	GACCGTGGTT	TCTGTGTGGT	GTCACCTGCC	AGACCCGCCG	AAGAAGCCAC
901	CTCTTGGAG	GGTGCCTCT	CTGGCACCGC	CCACTCCCAC	CCATCCGTGG
951	GCCGCCAGCA	CCACGGGGC	CCCCATCCA	CATCGCGGCC	ACCACGTCC
1001	GGGACACGCC	TTGTCCCCCG	GTGTACGCCG	AGACCAAGCA	CTTCCTCTAC
1051	TCCTCAGGCG	ACAAGNACAC	TGCGNCCTC	CTTCCTACTC	AATATATCTG
1101	AGGCCAGCC	TGACTGGCGT	TCGGGAGGTT	CGTGGAGACA	NTCTTCTGG
1151	TTCCAGGCCT	TGGATGCCAG	GATTCCCCGC	AGGTTGCC	GCCTGCCCA
1201	GCGNTACTGG	CAAATGCGGC	CCCTGTTCT	GGAGCTGCTT	GGGAACCACG
1251	CGCAGTGCCC	CTACGGGGTG	TTCCTCAAGA	CGCACTGCC	GCTGCGAGCT
1301	GCGGTCAACC	CAGCAGCCGG	TGTCTGTGCC	CGGGAGAAAGC	CCCAGGGCTC
1351	TGTGGCGGCC	CCCGAGGAGG	AGGAACACAG	ACCCCCGTCC	CCTGGTGCAG
1401	CTGCTCCGCC	AGCACACAG	CCCCGGCAG	GTGTACGGCT	TGTCGCGGGC
1451	CTGCTCGCG	CGGCTGGTGC	CCCCAGGGCT	CTGGGGCTCC	AGGCACAACG
1501	AACCCGCTT	CCTCAGGAAC	ACCAAGAAGT	TCATCTCCCT	GGGAAAGCAT
1551	GCCAAGCTCT	CGCTGCAGGA	GCTGACGTGG	AAGATGAGCG	TGCGGGACTG
1601	CGCTTGGCTG	CGCAGGAGCC	CAGGGGTTGG	CTGTGTTCCG	CCGCAGAGGC
1651	ACCGTCTGCG	TGAGGAGATC	CTGGCCAAGT	TCCTGCACTG	GCTGATGAGT
1701	GTGTACGTG	TCGAGCTGCT	CAGGTCTTTC	TTTTATGTCA	CGGAGACAC
1751	GTTCAAAAG	AACAGGCTCT	TTTTCTACCG	GAAGAGTGTG	TGGAGCAAGT
1801	TGCAAAGCAT	TGGAATCAGA	CAGCACTTGA	AGAGGGTGCA	GCTGCGGGAG
1851	CTGTCGGAAG	CAGAGGTCA	GCAGCATCGG	GAAGCCAGGC	CCGCCCTGCT
1901	GACGTCCAGA	CTCCGTTCA	TCCCCAAGCC	TGACGGGCTG	CGGCCGATTG
1951	TGAACATGGA	CTACGTGCG	GGAGCCAGAA	CGTCCGCAG	AGAAAAGAGG
2001	GCCGAGCGTC	TCACCTCGAG	GGTGAAGGC	CTGTTCAGCG	TGCTCAACTA
2051	CGAGCGGGCG	CGGCGCCCG	GCCTCCTGGG	CGCCTCTGTG	CTGGGCCTGG
2101	ACGATATCCA	CAGGGCCTGG	CGCACCTCG	TGCTGCGTGT	CGGGGCCAG
2151	GACCCGCGC	CTGAGCTGTA	CTTTGTCAAG	GTGGATGTGA	CGGGCGCGTA
2201	CGACACCATC	CCCCAGGACA	GGCTCACCGA	GGTCATCGCC	AGCATCATCA
2251	AACCCAGAA	CACGTACTGC	GTGCGTGGT	ATGCCGTGGT	CCAGAAGGCC
2301	GCCCATGGGC	ACGTCCGCAA	GGCCTTCAAG	AGCCACGTCT	CTACCTTGAC
2351	AGACCTCCAG	CCGTACATGC	GACAGTTCGT	GGCTCACCTG	CAGGANAACA
2401	GCCCGCTGAG	GGATGCCGTC	GTCATCGAGC	AGAGCTCCTC	CCTGAATGAG
2451	GCCAGCAGTG	GCCTCTCGA	CGTCTTCCTA	CGCTTCATGT	GCCACACAGC

FIG. 50A



+

62/78

2501	CGTGCATC	AGGGCAAGT	CCTACGTC	GTGCCAGGG	ATCCGCAGG
2551	GCTCCATCCT	CTCCACGCTG	CTCTGCAGCC	TGTGCTACGG	CGACATGGAG
2601	AACAAGCTGT	TTGCGGGGAT	TCGGCGGGAC	GGGCTGCTCC	TGCGTTGGT
2651	GGATGATTTC	TTGTTGGTGA	CACCTCACCT	CACCCACGCG	AAAACCTTC
2701	TCAGGACCT	GGTCCGAGGT	GTCCCCTGAGT	ATGGCTGCGT	GGTGAACTTG
2751	CGGAAGACAG	TGGTGAACCT	CCCTGTAGAA	GACGAGGCC	TGGGTGGCAC
2801	GGCTTTGTT	CAGATGCCGG	CCCACGGCCT	ATTCCCCTGG	TGCGGCCTGC
2851	TGCTGGATAC	CCGGACCCTG	GAGGTGCAGA	GCGACTACTC	CAGCTATGCC
2901	CGGACCTCCA	TCAGAGCCAG	TCTCACCTC	AACCGCGGCT	TCAAGGCTGG
2951	GAGGAACATG	CGTCGAAAC	TCTTTGGGGT	CTTGCGGCTG	AAGTGTACAA
3001	GCCTGTTCT	GGATTGCGAG	GTGAACAGCC	TCCAGACGGT	GTGCACCAAC
3051	ATCTACAAGA	TCCTCCTGCT	GCAGGGTAC	AGGTTTCACG	CATGTGTGCT
3101	GCAGCTCCA	TTTCATCAGC	AAAGTTGGAA	GAACCCACAA	TTTTCTCTGC
3151	CGCTCATCTC	TGACACGGCC	TCCCTCTGCT	ACTCCATCCT	GAAAGCCAAG
3201	AACCGAGGGA	TGTCGCTGGG	GGCCAAGGGC	GCCGCCGGCC	CTCTGCCCTC
3251	CGAGGCCGTG	CAGTGGCTGT	GCCACCAAAGC	ATTCCCTGCTC	AAGCTGACTC
3301	GACACCGTGT	CACCTACGTG	CCACTCCTGG	GGTCACTCAG	GACAGCCCAG
3351	ACGCAGCTGA	GTCGGAAGCT	CCCGGGGACG	ACGCTGACTG	CCCTGGAGGC
3401	CGCAGCCAAC	CCGGCACTGC	CCTCAGACTT	CAAGACCATC	CTGGACTGAT
3451	GGCCACCCGC	CCACAGCCAG	GCCGAGAGCA	GACACCAGCA	GCCCTGTCAC
3501	GCCCCGCTCT	ACGTCCCAGG	GAGGGAGGGG	CGGCCACAC	CCAGGCCCGC
3551	ACCGCTGGGA	GTCTGAGGCC	TGAGTGAGTG	TTTGGCCGAG	GCCTGCATGT
3601	CCGGCTGAAG	GCTGAGTGTC	CGGCTGAGGC	CTGAGCGAGT	GTCCAGCAA
3651	GGGCTGAGTG	TCCAGCACAC	CTGCCGTCTT	CACTTCCCCA	CAGGCTGGCG
3701	CTCGGCTCCA	CCCCAGGGCC	AGCTTTCT	CACCAAGGAGC	CCGGCTTCCA
3751	CTCCCCACAT	AGGAATAGTC	CATCCCCAGA	TTCGCCATTG	TTCACCCCTC
3801	GCCCTGCCCT	CCTTGCCCT	CCACCCCCAC	CATCCAGGTG	GAGACCCCTGA
3851	GAAGGACCT	GGGAGCTCTG	GGAAATTGGA	GTGACCAAAG	GTGTGCCCTG
3901	TACACAGGCG	AGGACCTGC	ACCTGGATGG	GGGTCCCTGT	GGGTCAAATT
3951	GGGGGGAGGT	GCTGTGGGAG	AAAAATACTG	AATATATGAG	TTTTTCAGTT
4001	TTGAAAAAAA	AAAAAAA	AAAAAAA	AAAAAAA	AAAAAAA

FIG. 50B

+



63/78

GCAGCGCTGCGTCCGTGCGCACGTGGAAAGCCCTGGCCCCGGCCACCCCGCGATGCC
 1 CGTCGCGACGCAGGACGACCGTGACCCCTCGGGACCGGGCGGTGGGGCGCTACGG 60

 a A A L R P A A H V G S P G P G H P R D A -
 b Q R C V L L R T W E A L A P A T P A M P -
 c S A A S C C A R G K P W P R P P P R C R -

 GCGCGCTCCCCGCTGCCGAGCCGTGCCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCT
 61 CGCGCGAGGGCGACGGCTCGGACGGACGCGTCGGTATGGCGCTCCACGA 120

 a A R S P L P S R A L P A A Q P L P R G A -
 b R A P R C R A V R S L L R S H Y R E V L -
 c A L P A A E P C A P C C A A T T A R C C -

 GCCGCTGGCCACGTTCGTGC GGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGG
 121 CGGCACCGGTGCAAGCACGCCGGACCCGGGTCCCACCGCCGACCACGTCGCC 180

 a A A G H V R A A P G A P G L A A A G A A R -
 b P L A T F V R R L G P Q G W R L V Q R G -
 c R W P R S C G A W G P R A G G W C S A G -

 GGACCCGGGGCTTCGGCGNTGGTGGCCANTGCNTGGTGTGGCTGCCCTGGANGN
 181 CCTGGGCCGCAAAGGCGCGNACCACCGGTNACGNACCACACGCACGGGACCCCTNCN 240

 a G P G G F P R ? G G P ? ? G V R A L G ? -
 b D P A A F R A ? V A ? C ? V C V P W ? ? -
 c T R R L S A R W W P ? A W C A C P G ? ? -

 ANGGCNGCCCCCGCCGCCCTCCCTCCGCCAGGTGTCCCTGCCTGAANGANCTGGTGGC
 241 TNCCGNCGGGGGCGGCCGGAGGAAGGCGGTCCACAGGACGGACTNCTNGACCACCG 300

 a ? A A P R R P L L P P G V L P E ? ? G G -
 b ? ? P P A A P S F R Q V S C L ? ? L V A -
 c G ? P P P P P P S A R C P A * ? ? W W P -

 CCGAGTGTGCANANGCTGTGCGANCAGCGCGCGAANAACGTGCTGGCCTTCGGCTTCGC
 301 GGCTCACGACGTNTNCGACACGCTNGCGCCGCGCTTNTGCACGACCGGAAGCCGAAGCG 360

 a P S A A ? A V R ? R R E ? R A G L R L R -
 b R V L ? ? L C ? R G A ? N V L A F G F A -
 c E C C ? ? C A ? A A R ? T C W P S A S R -

 GCTGCTGGACGGGGCCCGCGGGGCCCCCGAGGCCTTCACCACCGCGCAGCTA
 361 CGACGACCTGCCCCGGCGCCCCGGGGCTCCGGAAAGTGGTGGTCGACCGCGTCGAT 420

 a A A G R G P R G P P R G L H H Q R A Q L -
 b L L D G A R G G P P E A F T T S V R S Y -
 c C W T G P A G A P P R P S P P A C A A T -

 CCTGCCAACACGGTGACCGACGCACTGCGGGGAGCGGGCGTGGGGCTGCTGCG
 421 GGACGGGTTGTGCCACTGGCTGCGTACGCCCGCACCCCGACGACGACGC 480

 a P A Q H G D R R T A G E R G V G A A A A -
 b L P N T V T D A L R G S G A W G L L L R -
 c C P T R * P T H C G G A G R G G C C C A A -

FIG. 51A



64/78

a P R G R R R A G S P A G T L R ? ? C A G -
 b R V G D D V L V H L L A R C A ? F V L V -
 c A W A T T C W F T C W H A A R ? L C W W -

 541 GGNNTCCCAGCTGCGCCTACCAANGTGTGCGGGCCGCCGCTGTACCAAGCTCGGCCTGCNAC
 CCNAGGGTCGACGCCGATGGTNCACACGCCGCCGACATGGTCGAGCCGACGNTG + 600

 a G S Q L R L P ? V R A A A V P A R R C ? -
 b ? P S C A Y ? V C G P P L Y Q L G A A T -
 c ? P A A P T ? C A G R R C T S S A L ? L -

 601 TCAGGCCCGGCCCCGCCACACGCTANTGGACCCAANGCGTCTGGGATCCAACGGCCT
 AGTCCGGGCCGGGGCGGTGTGCGATNACCTGGCTTNCGCAGACCCTAGGTTGCCCGGA + 660

 a S G P A P A T R ? W T R ? R L G S N G P -
 b Q A R P P P H A ? G P E ? V W D P T G L -
 c R P G P R H T L ? D P ? A S G I Q R A W -

 661 GGAACCATAAGCGTCAGGGAGGCCGGGGCTCCCCCTGGCTGCCAGCCCCGGTGCAGGGAG + 720
 CCTTGGTATCGCAGTCCTCCGGCCCCAGGGGACCCGACGGTCGGGCCACGCTCCTC

 a G T I A S G R P G S P W A A S P G C E E -
 b E P * R Q G G R G P P G L P A P G A R R -
 c N H S V R E A G V P L G C Q P R V R G G -

 721 GCGCGGGGGCAGTGCCAGGCCAAGTCTGCCGTTGCCAAGAGGCCAGGGTGGCGCTGC
 CGCGCCCCCGTCACGGTCGGCTTCAGACGGCAACGGTTCTCCGGTCCGCACCGCGACG + 780

 a A R G Q C Q P K S A V A Q E A A Q A W R C -
 b R G G S A S R S L P L P K R P R R G A A -
 c A G A V P A E V C R C P R G P G V A L P -

 781 CCCTGAGCCGGAGCGGACGCCGTTGGCAGGGGCTGGGCCACCCGGCAGGACGCC
 GGGACTCGGCCTCGCCTGCCGGCAACCCGTCCCCAGGACCCGGTGGGCCCTCGCG + 840

 a P * A G A D A R W A G V L G P P G Q D A -
 b P E P E R T P V G Q G S W A H P G R T P -
 c L S R S G R P L G R G P G P T R A G R L -

 841 TGGACCGAGTGACCGTGGTTCTGTGTGGTGTACCTGCCAGACCCGCCAAGAACCC
 ACCTGGCTACTGGCACAAAGACACACCAGTGGACGGTCTGGCGGCTTCGGTG + 900

 a W T E * P W F L C G V T C Q T R R R S H -
 b G P S D R G F C V V S P A R P A E E A T -
 c D R V T V V S V W C H L P D P P K K P P -

 901 CTCTTGAGGGTGCCTCTGGCACGCCACTCCACCCATCCGTGGCCGCCAGCA
 GAGAAACCTCCCACGCCAGAGACCGTGCAGGGTGGTAGGCACCCGGCGGTG + 960

 a L F G G C A L W H A P L P P I R G P P A -
 b S L E G A L S G T R H S H P S V G R Q H -
 c L W R V R S L A R A T P T H P W A A S T -

 961 CCACCGGGCCCCCATCCACATCGCCGCCACACGTCCCTGGGACAGCCTTGTCCCCCG
 GGTGGCCCGGGGGTAGGTGTAGCGCCGGTGGTGCAGGACCCGTGCGGAACAGGGG + 1020

FIG. 51B

+



65/78

a P R G P P I H I A A T T S W D T P C P P -
 b H A G P P S T S R P P R P G T R L V P R -
 c T R A P H P H R G H H V L G H A L S P G -

 GTGTACGCCAGACCAAGCACTTCCTACTCCTCAGGCAGACAGNACACTGCGNCCCTC
 1021 -----+-----+-----+-----+-----+ 1080
 CACATGCGGCTCTGGTTCGTGAAGGAGATGAGGAGTCCGCTGTTCTGTGACGCNGGGAG

 a V Y A E T K H F L Y S S G D K ? T A ? L -
 b C T P R P S T S S T P Q A T ? T L R P S -
 c V R R D Q A L P L L L R R Q ? H C ? P P -

 CTTCTACTCAATATATCTGAGGCCAGCCTGACTGGCCTCGGGAGGTCGTGGAGACA
 1081 -----+-----+-----+-----+-----+ 1140
 GAAGGATGAGTTATATAGACTCCGGTCCGACTGACCGAACGCCCTCCAAGCACCTCTGT

 a L P T Q Y I * G P A * L A F G R F V E T -
 b F L L N I S E A Q P D W R S G G S W R ? -
 c S Y S I Y L R P S L T G V R E V R G D ? -

 NTCTTTCTGGTTCAGGCCCTGGATGCCAGGAITCCCCGCAGGTTGCCCGCCTGCCCA
 1141 -----+-----+-----+-----+-----+ 1200
 NAGAAAGACCAAGGTCCGAACTACGGTCTAAGGGCGTCCAACGGGCGGACGGGT

 a ? F L V P G L G C Q D S P Q V A P P A P -
 b S F W F Q A L D A R I P R R L P R L P Q -
 c L S G S R P W M P G F P A G C P A C P S -

 GCGNTACTGGCAAATGCCGCCCTGTTCTGGAGCTGCTTGGGAACCAACGCCAGTGCC
 1201 -----+-----+-----+-----+-----+ 1260
 CGCNATGACCGTTACGCCGGGACAAAGACCTCGACGAACCTTGGTCCGCGTCACGGG

 a A ? L A N A A P V S G A A W E P R A V P -
 b R Y W Q M R P L F L E L L G N H A Q C P -
 c ? T G K C G P C F W S C L G T T R S A P -

 CTACGGGTGTTCTCAAGACGCACTGCCGCTGCGAGCTGCGGTACCCCAGCAGCCGG
 1261 -----+-----+-----+-----+-----+ 1320
 GATGCCCAACAGGAGTTCTGCGTGACGGCGACGCTCGACGCCAGTGGGTCGTCGCC

 a L R G V P Q D A L P A A S C G H P S S R -
 b Y G V F L K T H C P L R A A V T P A A G -
 c T G C S S R R T A R C E L R S P Q Q P V -

 TGCTCTGCCGGGAGAACGCCAGGGCTCTGTGCGGCCCGAGGAGGAACACAG
 1321 -----+-----+-----+-----+-----+ 1380
 ACAGACACGGGCCCTTCGGGTCCGAGACACCGCCGGGCTCTCCTTGTGTC

 a C L C P G E A P G L C G G P R G G G T Q -
 b V C A R E K P Q G S V A A P E E E E H R -
 c S V P G R S P R A L W R P P R R R N T D -

 ACCCCCGTCGCTGGTGCAGCTGCTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGGCT
 1381 -----+-----+-----+-----+-----+ 1440
 TGGGGCAGCGGACCGTCGACGAGGCGGTCTGTCGTCGGGACCGTCCACATGCCGA

 a T P V A W C S C S A S T A A P G R C T A -
 b P P S P G A A A P P A Q Q P L A G V R L -
 c P R R L V Q L L R Q H S S P W Q V Y G F -

 TCGTGCAGGCCCTGCCTGCGCCGGCTGGTGCCTCTGGGCTCCAGGCACAAAG
 1441 -----+-----+-----+-----+-----+ 1500
 AGCACGCCCGGACGGACGCCAGGGGGTCCGGAGACCCCGAGGTCCGTGTTGC

FIG. 51C

+



66/78

S C G P A C A G W C P Q A S G A P G T T
 R A G L P A P A G A P R P L G L Q A Q R
 V R A C L R R L V P P G L W G S R H N E

 AACGCCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAAGCATGCCAAGCTCT
 1501 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTGCGGCGAAGGAGTCCTTGTTCTCAAGTAGAGGGACCCCTCGTACGGTTCGAGA

 N A A S S G T P R S S S P W G S M P S S -
 T P L P Q E H Q E V H L P G E A C Q A L -
 R R F L R N T K K F I S L G K H A K L S -

 CGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCAGGACTGCGCTGGCTGCGCAGGAGCC
 1561 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 GCGACGTCCTCGACTGCACCTCTACTCGCACGCCCTGACGCGAACGACGCCGTCCTCGG

 R C R S * R G R * A C G T A L G C A G A -
 A A G A D V E D E R A G L R L A A Q E P -
 L Q E L T W K M S V R D C A W L R R S P -

 CAGGGGTTGGCTGTGTTCCGGCCGCAGAGCACCGTCTGCGTGAGGAGATCCTGGCCAAGT
 1621 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 GTCCCCAACCGACACAAGGCCGGCGTCTGAGCAGCCTCTAGGACCAGGTTCA

 Q G L A V F R P Q S T V C V R R S W P S -
 R G W L C S G R R A P S A * G D P G Q V -
 G V G C V P A A E H R L R E E I L A K F -

 TCCTGCACTGGCTGATGAGTGTGTACGTCGTCAGCTGCTCAGGTCTTTCTTTATGTCA
 1681 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 AGGACGTGACCGACTACTCACACATGCAGCAGCTGACGAGTCCAGAAAGAAAATACAGT

 S C T G * * V C T S S S C S G L S F M S -
 P A L A D E C V R R R A A Q V F L L C H -
 L H W L M S V Y V V E L L R S F F Y V T -

 CGGAGACCACGTTCAAAAGAACAGGGCTTTCTACCGGAAGAGTGTCTGGAGCAAGT
 1741 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 GCCTCTGGTGCAAAGTTCTGTCGAGAAAAAGATGGCCTCTCACAGACCTCGITCA

 R R P R F K R T G S F S T G R V S G A S -
 G D H V S K E Q A L F L P E E C L E Q V -
 E T T F Q K N R L F F Y R K S V W S K L -

 TGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCTGCGGAAG
 1801 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 ACGTTTCGTAACCTTAGTCTGTCGTAACCTCTCCACGTCGACGCCCTCGACAGCCTTC

 C K A L E S D S T * R G C S C G S C R K -
 A K H W N Q T A L E E G A A A G A V G S -
 Q S I G I R Q H L K R V Q L R E L S E A -

 CAGAGGTCAGGCAGCATCGGAAGCCAGGCCGCCCTGCTGACGTCCAGACTCCGTTCA
 1861 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 GTCTCCAGTCGTCGTAGCCCTCGGTCCGGAGCAGACTGCAGGTCTGAGGCGAAGT

 Q R S G S I G K P G P P C * R P D S A S -
 R G Q A A S G S Q A R P A D V Q T P L H -
 E V R Q H R E A R P A L L T S R L R F I -

 TCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTTGGAGCCAGAA
 1921 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 AGGGGTTCGGACTGCCCCACGCCGGCTAACACCTGTACCTGATGCAAGCACCCTCGGTCTT

FIG. 51D



67/78

FIG. 51E



68/78

a	A R * G M P S S S S R A P P * M R P A V -
b	P A E G C R R H R A E L L P E * G Q Q W -
c	P L R D A V V I E Q S S S L N E A S S G -
	GCCTCTTCGACGTCTCCTACGTTCATGTGCCACCACGCCGTGCGCATCAGGGCAAGT
2461	-----+-----+-----+-----+-----+-----+-----+ 2520
	CGGAGAAGCTGAGAAGGATGCGAAGTACACGGTGGTGGCACCGTAGTCCCCGTCA
a	A S S T S S Y A S C A T T P C A S G A S -
b	P L R R L P T L H V P P R R A H Q G Q V -
c	L F D V F L R F M C H H A V R I R G K S -
	CCTACGTCCAGTGCAGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTGCAGCC
2521	-----+-----+-----+-----+-----+-----+-----+ 2580
	GGATGCAGGTACGGTCCCCTAGGGCGTCCCGAGGTAGGAGAGGTGCGACGAGACGTCGG
a	P T S S A R G S R R A P S S P R C S A A -
b	L R P V P G D P A G L H P L H A A L Q P -
c	Y V Q C Q G I P Q G S I L S T L L C S L -
	TGTGCTACGGGACATGGAGAACAAAGCTGTTGGGGATTCCGGGGACGGGCTGCTCC
2581	-----+-----+-----+-----+-----+-----+-----+ 2640
	ACACGATGCCGCTGTACCTCTTGTGACAAACGCCCTAACGCCGCCCCGACGAGG
a	C A T A T W R T S C L R G F G G T G C S -
b	V L R R H G E Q A V C G D S A G R A A P -
c	C Y G D M E N K L F A G I R R D G L L L -
	TGCGTTGGTGGATGATTCTTGTGACACCTCACCTCACCCACGCGAAAACCTTCC
2641	-----+-----+-----+-----+-----+-----+-----+ 2700
	ACGCAAACCACTACTAAAGAACAAACCACTGTGGAGTGGAGTGGGTGCGCTTTGGAGG
a	C V W W M I S C W * H L T S P T R K P S -
b	A F G G * F L V G D T S P H P R E N L P -
c	R L V D D F L L V T P H L T H A K T F L -
	TCAGGACCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTAACCTGCGGAAGACAG
2701	-----+-----+-----+-----+-----+-----+-----+ 2760
	AGTCCTGGGACCAGGCTCACAGGACTCATACCGACGCACCACITGAACGCCCTGTGTC
a	S G P W S E V S L S M A A W * T C G R Q -
b	Q D P G P R C P * V W L R G E L A E D S -
c	R T L V R G V P E Y G C V V N L R K T V -
	TGGTGAACCTCCCTGTAGAACAGCAGGGCCCTGGTGGCACGGCTTTGTTAGATGCCGG
2761	-----+-----+-----+-----+-----+-----+-----+ 2820
	ACCACTTGAGGGACATCTCTGCTCCGGGACCCACCGTGGCAAACAAAGTCTACGCC
a	W * T S L * K T R P W V A R L L F R C R -
b	G E L P C R R R G P G W H G F C S D A G -
c	V N F P V E D E A L G G T A F V Q M P A -
	CCCACGGCTATTCCCTGGTGCAGGCGCTGCTGGATAACCGGACCCCTGGAGGTGCAGA
2821	-----+-----+-----+-----+-----+-----+-----+ 2880
	GGGTGCCGATAAGGGACACGCCGGACGACGACCTATGGGCTGGGACCTCACGTCT
a	P T A Y S P G A A C C C W I P G P W R C R -
b	P R P I P L V R P A A G Y P D P G G A E -
c	H G L F P W C G L L L D T R T L E V Q S -
	GCGACTACTCCAGCTATGCCGGACCTCCATCAGAGCCAGTCTCACCTCAACCGCGCT
2881	-----+-----+-----+-----+-----+-----+-----+ 2940
	CGCTGATGAGGTGATAACGGGCTGGAGGTAGTCTCGGTCAAGAGTGGAAAGTTGGCGCGA

FIG. 51F

+



69/78

FIG. 51G



70/78

a	V G S S R G R R *	L P W R P Q P T R H C -
b	S E A P G D D A D C P G G R S Q P G T A -	
c	R K L P G T T L T A L E A A A N P A L P -	
	CCTCAGACTTCAGACCATCCTGGACTGATGCCACCGCCCACAGCCAGGCCAGAGCA	
3421	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3480 GGAGTCTGAAGTTCTGGTAGGACCTGACTACCGGTGGCGGGTGTGGTCCGGCTCTCGT	
a	P Q T S R P S W T D G H P P T A R P R A -	
b	L R L Q D H P G L M A T R P Q P G R E Q -	
c	S D F K T I L D * W P P A H S Q A E S R -	
	GACACCAGCAGCCCTGTCACGCCGGCTCTACGTCCCAGGGAGGGAGGGCGGCCACAC	
3481	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3540 CTGTGGTCGTGGGACAGTGCAGGGCCGAGATGCAGGGTCCCTCCCTCCCCGCCGGTGTG	
a	D T S S P V T P G S T S Q G G R G G P H -	
b	T P A A L S R R A L R P R E G G A A H T -	
c	H Q Q P C H A G L Y V P G R E G R P T P -	
	CCAGGCCCGCACCGCTGGGAGTCTGAGGCTTGAGTGAGTGTGTTGGCCGAGGGCTGATGT	
3541	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3600 GGTCCGGCGTGGCGACCCCTCAGACTCCGGACTCACTCACAAACCGCTCCGGACGTACA	
a	P G P H R W E S E A * V S V W P R P A C -	
b	Q A R T A G S L R P E * V F G R G L H V -	
c	R P A P L G V * G L S E C L A E A C M S -	
	CCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCCTGAGCGAGTGTCAGCCAAGGGCTGAGTG	
3601	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3660 GGCCGACTTCCGACTCACAGGCCGACTCCGGACTCGTCACAGGTGGTCCCGACTCAC	
a	P A E G * V S G * G L S E C P A K G * V -	
b	R L K A E C P A E A * A S V Q P R A E C -	
c	G * R L S V R L R P E R V S S Q G L S V -	
	TCCAGCACACCTGCCGTCTCACTTCCCCACAGGCTGGCGCTGGCTCCACCCCCAGGGCC	
3661	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3720 AGGTCTGTGGACGGCAGAAGTGAAGGGGTGTCCGACCGCGAGCCGAGGGTGGGTCCCGG	
a	S S T P A V F T S P Q A G A R L H P R A -	
b	P A H L P S S L P H R L A L G S T P G P -	
c	Q H T C R L H F P T G W R S A P P Q G Q -	
	AGCTTTCTCACCAAGGAGGCCGGCTTCACTCCCCACATAGGAATAGTCATCCCCAGA	
3721	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3780 TCGAAAAGGAGTGGTCCTGGGCCGAAGGTGAGGGGTGTATCCTTATCAGGTAGGGTCT	
a	S F S S P G A R L P L P T * E * S I P R -	
b	A F P H Q E P G F H S P H R N S P S P D -	
c	L F L T R S P A S T P H I G I V H P Q I -	
	TTCGCCATTGTCACCCCTGCCCTGCCCTCCACCCCCACCATCCAGGTG	
3781	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3840 AAGCGGTAACAAGTGGGAGCGGGACGGGAGGAACGGAAGGTGGGGTGGTAGGTCCAC	
a	F A I V H P S P C P P L P S T P T I Q V -	
b	S P L F T P R P A L L C L P P P P S R W -	
c	R H C S P L A L P S F A F H P H H P G G -	
	GAGACCTGAGAAGGACCCCTGGGAGCTGGGAATTGGAGTGACCAAAGGTGTGCCCTG	
3841	-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3900 CTCTGGGACTCTCCTGGACCCCTCGAGACCCCTAACCTCACTGGTTCCACACGGGAC	

FIG. 51H



+

71/78

a E T L R R T L G A L G I W S D Q R C A L -
 b R P * E G P W E L W E F G V T K G V P C -
 c D P E K D P G S S G N L E * P K V C P V -

TACACAGGCGAGGGACCCCTGCACCTGGATGGGGTCCCTGTGGTCAAATGGGGGAGGT
 3901 -----+-----+-----+-----+-----+-----+ 3960
 ATGTGTCCGCTCCTGGACGTGGACCTACCCCCAGGGACACCCAGTTAACCCCCCTCCA

a Y T G E D P A P G W G S L W V K L G G G -
 b T Q A R T L H L D G G P C G S N W G E V -
 c H R R G P C T W M G V P V G Q I G G R C -

GCTGTGGGAGTAAAATACTGAATATATGAGTTTCAGTTTGAAAAA
 3961 -----+-----+-----+-----+-----+-----+ 4020
 CGACACCCTCATTTATGACTTATATACTCAAAAGTCAAAACTTTTTTTTTT

a A V G V K Y * I Y E F F S F E K K K K K K -
 b L W E * N T E Y M S F S V L K K K K K K K -
 c C G S K I L N I * V F Q F * K K K K K K K -

AAAAAAA
 4021 ----- 4029
 TTTTTTTTT

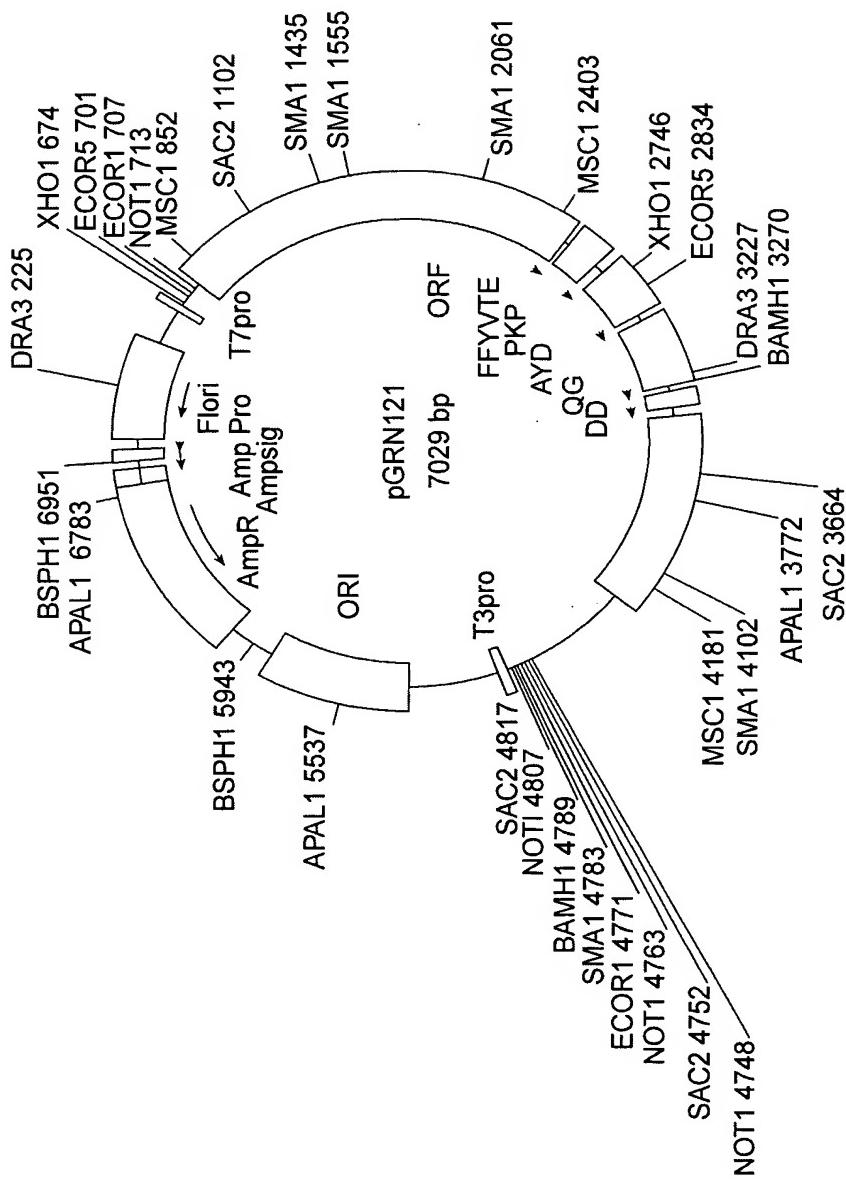
a K K K -
 b K K -
 c K K -

FIG. 51I

+



72/78



F/G. 52

+



73/78

1
met

GCAGCGCTGCGTCCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCG ATG

10

pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

20

his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

30

gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

40

phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

50

ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

60

leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

70

gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC CTG GCG CTG GAC GGG

80

ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC

90

tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

100

trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

110

leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

120

ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

130

thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

140

150

160

170

180

190

FIG. 53A



74/78

200

leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

220

val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

230

ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

240

ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

250

ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

260

val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

270

gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

280

gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

300

trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

310

leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

320

leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

330

glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

340

arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

350

leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

360

val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

370

380

390

400

410

420

FIG. 53B



75/78

430

ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

440

ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

460

leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

470

ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

490

his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

500

leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CAG GAG CTG ACG TGG AAG

520

met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

530

gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

550

ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

560

leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

580

arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC

590

ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

610

ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

620

leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

640

pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

FIG. 53C

+



76/78

650 660

arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

670

phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG CGC CCC GGC CTC CTG

680 690

gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

700

thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG

710 720

tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

730

gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

740 750

asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

760

his gly his val arg lys ala phe lys ser his val ser thr leu
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC TCT ACC TTG

770 780

thr asp leu gln pro tyr met arg gln phe val ala his leu gln
ACA GAC CTC CAG CCG TAC ATG CGA CAG TTC GTG GCT CAC CTG CAG

790

glu thr ser pro leu arg asp ala val val ile glu gln ser ser
GAG ACC AGC CCG CTG AGG GAT GCC GTC GTC ATC GAG CAG AGC TCC

800 810

ser leu asn glu ala ser ser gly leu phe asp val phe leu arg
TCC CTG AAT GAG GCC AGC AGT GGC CTC TTC GAC GTC TTC CTA CGC

820

phe met cys his his ala val arg ile arg gly lys ser tyr val
TTC ATG TGC CAC CAC GCC GTG CGC ATC AGG GGC AAG TCC TAC GTC

830 840

gln cys gln gly ile pro gln gly ser ile leu ser thr leu leu
CAG TGC CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC

850

cys ser leu cys tyr gly asp met glu asn lys leu phe ala gly
TGC AGC CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG

860 870

ile arg arg asp gly leu leu leu arg leu val asp asp phe leu
ATT CGG CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG

FIG. 53D

+



77/78

880

leu val thr pro his leu thr his ala lys thr phe leu arg thr
TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC

890

leu val arg gly val pro glu tyr gly cys val val asn leu arg 900
CTG GTC CGA GGT GTC CCT GAG TAT GGC TGC GTG AAC TTG CGG

910

lys thr val val asn phe pro val glu asp glu ala leu gly gly
AAG ACA GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC

920

thr ala phe val gln met pro ala his gly leu phe pro trp cys 930
ACG GCT TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC

940

gly leu leu leu asp thr arg thr leu glu val gln ser asp tyr
GGC CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC

950

ser ser tyr ala arg thr ser ile arg ala ser leu thr phe asn 960
TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC

970

arg gly phe lys ala gly arg asn met arg arg lys leu phe gly
CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG

980

val leu arg leu lys cys his ser leu phe leu asp leu gln val 990
GTC TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG

1000

asn ser leu gln thr val cys thr asn ile tyr lys ile leu leu
AAC AGC CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG

1010

leu gln ala tyr arg phe his ala cys val leu gln leu pro phe 1020
CTG CAG GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT

1030

his gln gln val trp lys asn pro thr phe phe leu arg val ile
CAT CAG CAA GTT TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC

1040

ser asp thr ala ser leu cys tyr ser ile leu lys ala lys asn 1050
TCT GAC ACG GCC TCC CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC

1060

ala gly met ser leu gly ala lys gly ala ala gly pro leu pro
GCA GGG ATG TCG CTG GGG GCC AAG GGC GCC GGC CCT CTG CCC

1070

ser glu ala val gln trp leu cys his gln ala phe leu leu lys 1080
TCC GAG GCC GTG CAG TGG CTG TGC CAC CAA GCA TTC CTG CTC AAG

1090

leu thr arg his arg val thr tyr val pro leu leu gly ser leu
CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC

FIG. 53E



78/78

1100

arg thr ala gln thr gln leu ser arg lys leu pro gly thr thr
AGG ACA GCC CAG ACG CAG CTG AGT CGG AAG CTC CCG GGG ACG ACG

1120

leu thr ala leu glu ala ala ala asn pro ala leu pro ser asp
CTG ACT GCC CTG GAG GCC GCA AAC CCG GCA CTG CCC TCA GAC

1130 1132

phe lys thr ile leu asp OP
TTC AAG ACC ATC CTG GAC TGA TGGCCACCCGCCACAGCCAGGCCAGAGCAGA

CACCAGCAGCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGCGGCCACACCC

AGGCCCGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTGTTGGCCGAGGCCTGCATGTCC

GGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTC

CAGCACACCTGCCGTCTCACTCCCCACAGGCTGGCGCTGGCTCCACCCCCAGGCCAG

CTTTCYTCACCAGGAGCCGGCTTCCACTCCCCACATAGGAATAGTCCATCCCCAGATT

CGCCATTGTTACCCYTCGCCCTGCCYTCCTTGCCCTCCACCCCCACCATCCAGGTGGA

GACCCTGAGAAGGACCCCTGGGAGCTCTGGAAATTGGAGTGACCAAAGGTGTGCCCTGTA

CACAGGCAGGACCCCTGCACCTGGATGGGGTCCCTGTGGTCAAATTGGGGGAGGTGC

TGTGGGAGTAAAATACTGAATATATGAGTTTCAGTTTGRAAAAAAAAAAAAAAAA

AAAAAAAAAA

FIG. 53F

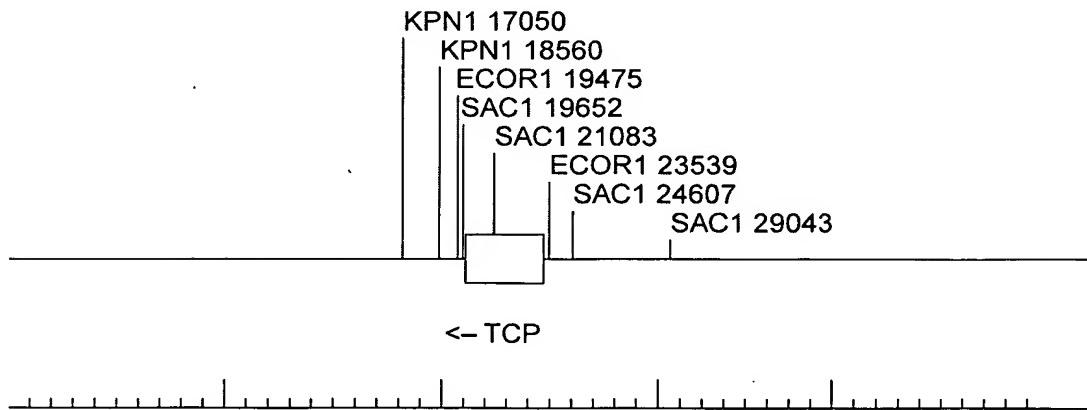


FIG. 54